

8

SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB#

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

BEST AVAILABLE COPY

STAFF USE ONLY
Contact: Sheppard

Searcher _____

Searcher Phone # tel: 308-4499

Searcher Location _____

Date Searcher Picked Up _____

Date Completed 10/19/01

Searcher Prep & Review Time _____

Clencal Prep Time _____

Online Time _____

Type of Search

NA Sequence (#) _____

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel Orbit _____

Dr Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 19:58:14 ; Search time 8321.72 Seconds
(without alignments)
3148.785 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTCGCGAGCCGCGACC.....TGTAAAAA 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	gb_est1:
2:	gb_est2:
3:	gb_est3:
4:	gb_est4:
5:	gb_est5:
6:	gb_est6:
7:	gb_est7:
8:	gb_est8:
9:	gb_est9:
10:	gb_est10:
11:	gb_est11:
12:	gb_est12:
13:	gb_est13:
14:	gb_est14:
15:	gb_est15:
16:	gb_est16:
17:	gb_est17:
18:	gb_est18:
19:	gb_est19:
20:	gb_est20:
21:	gb_est21:
22:	gb_est22:
23:	gb_est23:
24:	gb_est24:
25:	gb_est25:
26:	gb_est26:
27:	gb_est27:
28:	gb_est28:
29:	gb_est29:
30:	gb_est30:
31:	gb_est31:
32:	gb_est32:
33:	em_estba:
34:	em_estfun:
35:	em_esthum1:
36:	em_esthum2:
37:	em_esthum3:
38:	em_esthum4:
39:	em_esthum5:
40:	em_esthum6:
41:	em_esthum7:
42:	em_esthum8:
43:	em_esthum9:
44:	em_esthum10:
45:	em_esthum11:
46:	em_esthum12:
47:	em_esthum13:
48:	em_esthum14:
49:	em_esthum15:
50:	em_esthum16:
51:	em_esthum17:
52:	em_esthum18:
53:	em_esthum19:
54:	em_esthum20:
55:	em_esthum21:
56:	em_esthum22:
57:	em_esthum23:
58:	em_esthum24:
59:	em_esthum25:
60:	em_esthum26:
61:	em_esthum27:
62:	em_esthum28:
63:	em_estin1:
64:	em_estin2:
65:	em_estin3:
66:	em_estin4:
67:	em_estin5:
68:	em_estom1:
69:	em_estom2:
70:	em_estov1:
71:	em_estov2:
72:	em_estpl1:
73:	em_estpl2:
74:	em_estpl3:
75:	em_estpl4:
76:	em_estpl5:
77:	em_estpl6:
78:	em_estpl7:
79:	em_estpl8:
80:	em_estpl9:
81:	em_estpl10:
82:	em_estro1:
83:	em_estro2:
84:	em_estro3:
85:	em_estro4:
86:	em_estro5:
87:	em_estro6:
88:	em_estro7:
89:	em_estro8:
90:	em_estro9:
91:	em_estro10:
92:	em_estro11:
93:	em_estro12:
94:	em_estro13:
95:	em_estro14:
96:	em_estro15:
97:	em_estro16:
98:	em_estro17:
99:	em_estro18:
100:	em_estro19:
101:	em_estro20:
102:	gb_est25:
103:	gb_est26:
104:	gb_est27:
105:	gb_est28:
106:	gb_est29:
107:	gb_est30:
108:	gb_est31:
109:	gb_est32:
110:	gb_est41:
111:	gb_est42:
112:	gb_est43:
113:	gb_est44:
114:	gb_est45:
115:	gb_est46:
116:	gb_est47:

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

Source
1. .897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YGL2"
/clone_lib="LTI_FL011_BCI"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 152 a 291 c 268 g 182 t
ORIGIN

Query Match 30.5%; Score 844.8; DB 106; Length 897;
Best Local Similarity 98.5%; Pred. No. 7.8e-188;
Matches 879; Conservative 4; Mismatches 6; Indels 3; Gaps 3;

QY 80 GGGCAGGGCGCGCGCCACCGCCCGCGCATGGGGCTGAAGCGCGCCAGAGAGCGCT 139
DB 9 GCGCGAGGCGCGCGCCACCGCCCGCGCATGGGGCTGAAGCGCGCCAGAGAGCGCT 68
QY 140 GTTCCCGCTGCGCTCCATCGACGACCTGGTGGCCCTTTTGTGCGCGAGCTGGCGCCGAGA 199
DB 69 GTTCCCGCTGCGCTCCATCGACGACCTGGTGGCCCTTTTGTGCGCGAGCTGGCGCCGAGA 128
QY 200 GGAGCGGACCTGGTCTCTTTCCTTGGTGTGGGCTTCTGCGAGCATTTTCTGGCTGT 259
DB 129 GGAGCGGACCTGGTCTCTTTCCTTGGTGTGGGCTTCTGCGAGCATTTTCTGGCTGT 188
QY 260 CAACCGCTGATCCCTACCAACGCTTCCGAGCTACCTTCCAGCCCGCCAGCCCGCCCGCA 319
DB 189 CAACCGCTGATCCCTACCAACGCTTCCGAGCTACCTTCCAGCCCGCCAGCCCGCCCGCA 248
QY 320 CCGCGCTGGCGGCTCACCTACTTTCCCGTGGCGACCTGCTCTATCATCGCCGCCCTCTA 379
DB 249 CCGCGCTGGCGGCTCACCTACTTTCCCGTGGCGACCTGCTCTATCATCGCCGCCCTCTA 308
QY 380 TGCCCGCTTACCGCCCGAGATCCGAGCGCGCTGCGACCTGTCCCTCTATCTCGAGAAGG 439
DB 309 TGCCCGCTTACCGCCCGAGATCCGAGCGCGCTGCGACCTGTCCCTCTATCTCGAGAAGG 368
QY 440 GGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTATATGGAACACCTCAG 499
DB 369 GGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTATATGGAACACCTCAG 428
QY 500 CCCTCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCCAGCTTATCATCAGGAC 559
DB 429 CCCTCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCCAGCTTATCATCAGGAC 488
QY 560 CAAATTTGAGAGCTCCGGTGTGGCTTTTGTGTGGTGGGCGCTGCCAGCCCTTGGGTCT 619
DB 489 CAAATTTGAGAGCTCCGGTGTGGCTTTTGTGTGGTGGGCGCTGCCAGG-CCTGGGTCT 547
QY 620 CCGGATGTCCACCTCGCCCTGCTGAGGATCATGCTGGTGTAGTGTGGGCGCAATGG 679
DB 548 CCGGATGTCCACCTCGCCCTGCTGAGGATCATGCTGGTGTAGTGTGGGCGCAATGG 607
QY 680 GGAGCAGACAGCTGAGGTACCTGGCAGCGCAAGGGCAACGAGCGCCAGGGGCCAGAG 739
DB 608 GGAGCAGACAGCTGAGGTACCTGGCAGCGCAAGGGCAACGAGCGCCAGGGGCCAGAG 667

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	844.8	30.5	897	106	AL560097 AL560097
2	809.2	29.2	912	106	AL582775 AL582775
3	806.2	29.1	876	105	AL521543 AL521543
4	773.6	27.9	865	106	AL560096 AL560096
5	758.4	27.4	808	154	BG476481 BG476481
6	717.7	25.9	718	154	BG471808 BG471808
7	658.2	23.7	898	105	AL521542 AL521542
8	645.2	23.3	799	150	BF526328 BF526328
9	641.4	23.1	966	172	BF980770 BF980770
10	623.8	22.5	959	154	BG476263 BG476263
11	622.2	22.4	981	150	BF525363 BF525363
12	617.2	22.3	709	133	BG396934 BG396934
13	613.4	22.1	680	165	BE267789 BE267789
14	613	22.1	687	172	BF984524 BF984524
15	606	21.9	738	175	BG251389 BG251389
16	601.8	21.7	645	11	AA741057 AA741057
17	598.6	21.6	741	165	BE266998 BE266998
18	592.8	21.4	596	136	BE514777 BE514777
19	566.8	20.4	641	122	AW950638 AW950638
20	566.6	20.4	856	153	BG422364 BG422364
21	560.4	20.2	580	165	BE267618 BE267618
22	554.8	20.0	1139	169	BF795474 BF795474
23	551.6	19.9	764	137	BE573676 BE573676
24	547.6	19.8	898	146	BF237078 BF237078
25	545.4	19.7	547	165	BE274228 BE274228
26	540.6	19.5	630	165	BE281439 BE281439
27	535.4	19.3	687	152	BG326214 BG326214
28	532	19.2	574	122	AW968425 AW968425
29	531.6	19.2	692	143	BF012319 BF012319
30	516.4	18.6	624	114	AW321532 AW321532
31	505.4	18.2	792	106	AL561613 AL561613
32	488.4	17.6	507	112	AW150068 AW150068
33	484	17.5	523	169	BF803247 BF803247
34	477.4	17.2	479	105	AL079746 AL079746
35	466.8	16.8	525	11	AA776738 AA776738
36	466	16.8	533	11	AA743431 AA743431
37	461.8	16.7	525	19	AI333944 AI333944
38	460	16.6	532	136	BE532045 BE532045
39	459	16.6	503	15	AI082242 AI082242
40	449.4	16.2	497	104	AI953167 AI953167
41	448.8	16.2	496	116	AW439193 AW439193
42	445	16.1	490	152	BG328937 BG328937
43	444	16.0	511	10	AA705195 AA705195
44	443.8	16.0	635	136	BE533157 BE533157
45	442.8	16.0	457	191	Z98466 HS298466

ALIGNMENTS

RESULT 1
AL560097
LOCUS AL560097 LTI_FL011_BCI Homo sapiens cDNA clone CS0DG002YGL2 5 prime
DEFINITION AL560097 897 bp mRNA EST 16-FEB-2001
; mRNA sequence.

ACCESSION AL560097
VERSION AL560097.1 GI:12906230

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 897)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 153 a 281 c 256 g 180 t 6 others
ORIGIN

Query Match 29.1%; Score 806.2; DB 105; Length 876;
Best Local Similarity 96.4%; Pred. No. 9,1e-179;
Matches 845; Conservative 6; Mismatches 10; Indels 16; Gaps 2;

QY 90 CGCGGCCACGCGCCGATGGGCTGAAGCGGCCAGAACGCTGTCGCGCTG 149
Db 1 CGCGGCCACGCGCCGATGGGCTGAAGCGGCCAGAACGCTGTCGCGCTG 60
QY 150 CGCTCATCGACGCTGCTGGCGCTGTTGCTGCGAGCTGGCGGAGGAGCGGAC 209
Db 61 CGCTCATCGACGCTGCTGGCGCTGTTGCTGCGAGCTGGCGGAGGAGCGGAC 120
QY 210 CTGGTGCTCTTCCCTTGGTGTGGCTGCTGGAGCAATTTCTGGCTGTCAACCGCTC 269
Db 121 CTGGTGCTCTTCCCTTGGTGTGGCTGCTGGAGCAATTTCTGGCTGTCAACCGCTC 180
QY 270 ATCCCTACCAAGCTTCCGAGCTCACCCTTCAGCCGAGCCGCGCCGAGCGGCTGC 329
Db 181 ATCCCTACCAAGCTTCCGAGCTCACCCTTCAGCCGAGCCGCGCCGAGCGGCTGC 240
QY 330 GGCTCACCTACTTCCCTGCGGACCTGCTATCATCGCGCCCTCATGCGCGCTTC 389
Db 241 GGCTCACCTACTTCCCTGCGGACCTGCTATCATCGCGCCCTCATGCGCGCTTC 300
QY 390 ACCGCCAGATCCGAGGCGCCGTCGACCTGTCCCTCTATCTCGAAGGGGTGTCTCC 449
Db 301 ACCGCCAGATCCGAGGCGCCGTCGACCTGTCCCTCTATCTCGAAGGGGTGTCTCC 360
QY 450 AGCGGTGAGCTGTGAAGAGTCTCCGATGTCATATGAACAGCTCAGCGCTCTTAC 509
Db 361 AGCGGTGAGCTGTGAAGAGTCTCCGATGTCATATGAACAGCTCAGCGCTCTTAC 420
QY 510 TTCAAGGATCGGCGCCACATCCAGTCCCTCTCAGCTTCATCAC-----A 554
Db 421 TTCAAGGATCGGCGCCAAATCCAGTCCCTCTCAGCTTCATACAGGTTGGAGCCCAAGTA 480
QY 555 GGCACCAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGGCTTCGCCAGGCGCTG 614
Db 481 GGCACCAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGGCTTCGCCAGG-CTG 539
QY 615 GGTCTCCGGAGTTCACCTCGCCCTGTCTGAGGATCATGCTGGGTAGTGTGGGCC 674
Db 540 GGTCTCCGGAGTTCACCTCGCCCTGTATGAGGATCATGCTGGGTAGTGTGGGCC 599
QY 675 AATGGGAGCAGACAGCTGAGTCACTGGCAGCGGAAGGGCAACAGGACCGCAGGGC 734
Db 600 AATGGGAGCAGACAGTAGGTCACTGGCAGCGGAAGGGCAACAGGACCGCAGGGC 659

QY 735 CAGACAGTCAATCCGCTGCTGGCTGAGCGAGCTGGCTGTACCTGAAAGATCATACATG 794
Db 660 CAGACAGTCAATCCGCTGCTGGCTGAGCGAGTGGGCTGTACCTGAAAGATCATACATG 719
QY 795 CGCTGTGACCGCAAGATGAGGTGGGCTTCATGGTGTGCCATCAACCTTCCATTGAC 854
Db 720 CGCTGTGACCGCAAGATGAGGTGGGCTTCATGGTGTGCCATCAACCTTCCATTGAC 779
QY 855 CTGCACACGAGTCTGCTGGAGCTTCTGCAGCTGCACGAGAAGCTGCTTGGCTGCTCTAT 914
Db 780 CTGCACACGAGTCTGCTGGAGCTTCTGCAGCTGCACGAGAAGCTGCTTGGCTGCTCTAT 839
QY 915 GACCTGGGACATCTGGAAAGGTACCCCATGGCCTTAG 951
Db 840 GACCTGGGACATCTGGAAAGGTACCCCATGGCCTTAG 876

RESULT 4
AL560096/c 865 bp mRNA EST 16-FEB-2001
LOCUS AL560096 LTI_FL011_BCl Homo sapiens cDNA clone CS0DC002YG12 3 prime
DEFINITION mRNA sequence.
ACCESSION AL560096 GI:12906228
VERSION AL560096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DC002YG12"
/clone_lib="LTI_FL011_BCl"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 165 a 183 c 306 g 204 t 7 others
ORIGIN

Query Match 27.9%; Score 773.6; DB 106; Length 865;
Best Local Similarity 97.1%; Pred. No. 4,1e-171;
Matches 845; Conservative 5; Mismatches 14; Indels 6; Gaps 6;

QY 1700 AGTGGCAGCAGCGCTCAGTGGCCAGCACCAG-CAGCATCACCCAGCGGAGGTCCAG 1758
Db 865 AGTGGCAGCAGCGCTCAGTGGCCAGCACCAGCATCACCCAGCGGAGGTCCAG 806
QY 1759 TGCTCACTTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCTGCTGGTGGCCACCAAGA 1818
Db 805 TRCTCACTTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCTGCT-GTGGCCACCAAGA 747
QY 1819 TCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACAGTCGCAAGTGCAGATGAAGAAGC 1878

Db 746 TCAACTCGAGCGCCATCAAGTGCACACTCAGGCACAGTGCAGATGAAGAAGC 687
 Qy 1879 AGAAGTGTCCACCCCTAGTACACTCTGTCTTCTTCTCAAGCGCGCAAGGCC 1938
 Db 686 AGAAGTGTCCACCCCTAGTACACTCTGTCTTCTTCTCAAGCGCGCAAGGCC 627
 Qy 1939 TCTGAACCTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGCTTCCGCTTAGTCCGCCA 1998
 Db 626 TCTGAACCTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGCTTCCGCTTAGTCCGCCA 567
 Qy 1999 ACTCTAGCCCATGTTCTGCCCCCAGCCCCAAAGGGGACAGCGCTTACCTCTACCCAAACC 2058
 Db 566 ACTCTAGCCCATGTTCTGCCCCCAGCCCCAAAGGGGACAGCGCTTACCTCTACCCAAACC 507
 Qy 2059 CTAGTTTCCGGTCCCGAGTACAGTCTATCAACCCAGGATTTCTCAGCTCAGAAC 2118
 Db 506 CTAGTTTCCGGTCCCGAGTACAGTCTATCAACCCAGGATTTCTCAGCTCAGAAC 447
 Qy 2119 CCAGGGCTCTGCCCCAGTCTGTTAGATATAGTCTCTTCTCCAGAAATCCAGCGGCCA 2178
 Db 446 CCA-GGCTCTGCCCCAGTCTGTTAGATATAGTCTCTTCTCCAGAAATCCAGCGGCCA 388
 Qy 2179 ATGGAACCTACAGCTGGGTCTTAATACAGTCTTTAAAGGCCAGCGCTTAGAACCC 2238
 Db 387 ATGGAACCTACAGCTGGGTCTTAATACAGTCTTTAAAGGCCAGCGCTTAGAACCC 328
 Qy 2239 AAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTCTTACTCAGGGCTCTCTCCCA 2298
 Db 327 AAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTCTTACTCAGGGCTCTCTCCCA 268
 Qy 2299 GCTTGTAGGAGTGTAGTCTTACCTTTAACCCTTAACCCAAAGGGGACAGCGTCCCACTCCAGCCC 2358
 Db 267 GCTTGTAGGAGTGTAGTCTTACCTTTAACCCTTTAACCCTTAACCCAAAGGGGACAGCGTCCCACTCCAGCCC 209
 Qy 2359 GGGAGGCTAGGAGGACTCAGCCCTAGGAGTATATTTCCGCATTCAGAAATCCCATATC 2418
 Db 208 -GGAGGCTAGGAGGACTCAGCCCTAGGAGTATATTTCCGCATTCAGAAATCCCATATC 150
 Qy 2419 TTGGGAATCCAAGTCTCCCTGCCCCCAATAACTTCAGTCTTCCAGAAATTTGGAATC 2478
 Db 149 TTGGGAATCCAAGTCTCCCTGCCCCCAATAACTTCAGTCTTCCAGAAATTTGGAATC 91
 Qy 2479 CTAGTTTCTCTCTTCTGATATCCCGAGTGTGGGACACAAACTCCGCCCCAGCTATGA 2538
 Db 90 CTAGTTTCTCTCTTCTGATATCCCGAGTGTGGGACACAAACTCCGCCCCAGCTATGA 31
 Qy 2539 GCATCTGAGCGCCGCTCTTCTCTGACGA 2568
 Db 30 GCATCTGAGCGCCGCTCTTCTCTGACGA 1

RESULT 5
 BG476481 808 bp mRNA EST 21-MAR-2001
 LOCUS 60252015F1 NIH_MGC_20 Homo sapiens cDNA Clone IMAGE:4640450 5',
 DEFINITION mRNA sequence.
 ACCESSION BG476481
 VERSION BG476481.1 GI:13408760
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 808)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC).
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLC1405 row: n column: 03
 High quality sequence stop: 798.
 Location/Qualifiers
 source 1..808
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4640450"
 /lab_host="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 177 a 245 c 278 g 108 t
 ORIGIN

Query Match 27.4%; Score 758.4; DB 154; Length 808;
 Best Local Similarity 99.4%; Pred. No. 1.5e-167;
 Matches 803; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Qy 1092 CACTGTGCGAACCGCAATGTGCGGAAGCCCTGCAGCGCTGGCGGACACGGCCACTGTC 1151
 Db 2 CACTGTGCGAACCGCAATGTGCGGAAGCCCTGCAGCGCTGGCGGACACGGCCACTGTC 61
 Qy 1152 ATCCAGGACTACACTACTGCGGGGAAGACGAGGAGATCTACAGGAGTCTTTTGAAGTA 1211
 Db 62 ATCCAGGACTACACTACTGCGGGGAAGACGAGGAGATCTACAGGAGTCTTTTGAAGTA 121
 Qy 1212 GCCAATGATGTATCCCCAACCTGCTGGAAGGAGGACGACGCTTGTGTGAGGCGGCGAG 1271
 Db 122 GCCAATGATGTATCCCCAACCTGCTGGAAGGAGGACGACGCTTGTGTGAGGCGGCGAG 181
 Qy 1272 GAGGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCAAAGTTCCGCGCTCCAGGACCCCT 1331
 Db 182 GAGGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCAAAGTTCCGCGCTCCAGGACCCCT 241
 Qy 1332 GAGTGTCTGCCCCACCTGCTGCGATTCTACAGCGCATCTGCAATGGAGAGGCGAGT 1391
 Db 242 GAGTGTCTGCCCCACCTGCTGCGATTCTACAGCGCATCTGCAATGGAGAGGCGAGT 301
 Qy 1392 CCCACGCTGTGTGACAGTGGGCTGGGCGACCTTTCTTGTGAGTCCCTAGGCGCGTTT 1451
 Db 302 CCCACGCTGTGTGACAGTGGGCTGGGCGACCTTTCTTGTGAGTCCCTAGGCGCGTTT 361
 Qy 1452 GAGGACAGTGTGGGCAAGAGTTCGCATAGTACCGCAGAGCGCGGCGCGGAGGCC 1511
 Db 362 GAGGACAGTGTGGGCAAGAGTTCGCATAGTACCGCAGAGCGCGGCGCGGAGGCC 421
 Qy 1512 GAGGAGCGGTGGGCGAGGAAGCCCGGGAAGCGCGGCGGCGCGGCGGAGTCC 1571
 Db 422 GAGGAGCGGTGGGCGAGGAAGCCCGGGAAGCGCGGCGGCGGCGGAGTCC 481
 Qy 1572 AAGCCAGAGAGCGCGCGCGCCCAAGAGGACGACTGGACAGGCGCTGGGCGACCGGC 1631
 Db 482 AAGCCAGAGAGCGCGCGCGCCCAAGAGGACGACTGGACAGGCGCTGGGCGACCGGC 541
 Qy 1632 CAGGTCAGTGTGAGGACCGCGCGCGGGAAGCTCTCGGACTGTGCTGGCAGACCGCA 1691
 Db 542 CA-GGTGAGTGTGAGGACCGCGCGCGGGAAGCTCTCGGACTGTGCTGGCAGACCGCA 600
 Qy 1692 GGCCTGAAGTGTGAGCAGCGGCTCAGTGTGCGAGCAGCGGCGGAGTCAACCGCGGAG 1751
 Db 601 GGCCTGAAGTGTGAGCAGCGGCTCAGTGTGCGAGCAGCGGCGGAGTCAACCGCGGAG 660

```
QY 1752 GGTCCAGTGTCTTCCAGAGTGAAGAATGAAGGGCATG-AAGGAGCTGTCTGG-TGG 1809
Db 661 GTTCCAGTGTCTTCCAGAGTGAAGAATGAAGGGCATGAAAGGAGCTGTCTGGTTGG 720
QY 1810 CCACCAAGATCAACTCGAGCGGCATCAAGCTCAACTCAGGCACAG-TCGCAAGTGCAG 1868
Db 721 CCACCAAGATCAACTCGAGCGGCATCAAGCTCAACTCAGGCACAGTTCGCAAGTGCAG 780
QY 1869 ATGAAGAAGCAGAAAGTGTCCACCCCTTA 1896
Db 781 ATGAAGAAGCAGAAAGTGTCCACCCCTTA 808

RESULT 6
LOCUS BG471808 718 bp mRNA EST 21-MAR-2001
DEFINITION 602513561F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4636143 5',
mRNA sequence.
ACCESSION BG471808
VERSION BG471808.1 GI:13404083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 718)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC1394 row: j column: 16
High quality sequence stop: 718.
FEATURES
Location/Qualifiers
1..718
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4636143"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 155 a 216 c 216 g 131 t
ORIGIN
Query Match 25.9%; Score 717; DB 154; Length 718;
Best Local Similarity 100.0%; Pred. No. 7.9e-158;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 769 GCGTGTACCTGAAGGATCATACATCGCTGTGACCGCAGATGGAGGTGGCTTCATGG 828
Db 2 GCGTGTACCTGAAGGATCATACATCGCTGTGACCGCAGATGGAGGTGGCTTCATGG 61
QY 829 TGTGTGCATCAACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGC 888
Db 62 TGTGTGCCATCAACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGC 121
QY 889 AGCAGAAGCTGTCTGTGGCTGTATGACCTGGGACATCTGGAAGAGTACCCCATGGCCT 948
|||||
```

```
Db 122 AGCAGAAGCTGTCTGTGGCTGTCTATGACCTGGGACATCTGGAAGAGTACCCCATGGCCT 181
QY 949 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCCCAACCCCTGGCCCGCCAGACCCACTCA 1008
|||||
Db 182 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCCCAACCCCTGGCCCGCCAGACCCACTCA 241
QY 1009 CCTCTTACCACAAAGGCGATGTGCTCAGCCAAAGACCTTACTATCGGGATGAACACATCTACC 1068
Db 242 CCTCTTACCACAAAGGCGATGTGCTCAGCCAAAGACCTTACTATCGGGATGAACACATCTACC 301
QY 1069 CCTACATGTACCTGGCTGGCTACCACTGTGCAACCCGCAATGTGCGGGGAAGCCCTCGAG 1128
Db 302 CCTACATGTACCTGGCTGGCTACCACTGTGCAACCCGCAATGTGCGGGGAAGCCCTCGAG 361
QY 1129 CCTGGCGGCACACGGCACTGTCTATCCAGGACTTACAACTACTGCCGGGAAGACGAGAGA 1188
Db 362 CCTGGCGGCACACGGCACTGTCTATCCAGGACTTACAACTACTGCCGGGAAGACGAGAGA 421
QY 1189 TCTACAAGGAGTCTTTTGAAGTAGCCAAATGATGTCTATCCCAACCTTGTCTGAAGGAGCAG 1248
Db 422 TCTACAAGGAGTCTTTTGAAGTAGCCAAATGATGTCTATCCCAACCTTGTCTGAAGGAGCAG 481
QY 1249 CCAGCTTCTGGAGGCGGCGAGGAGCGCCGGGGGAGCAAAAGCCAGGACCCAGAGCC 1308
Db 482 CCAGCTTCTGGAGGCGGCGAGGAGCGCCGGGGGAGCAAAAGCCAGGACCCAGAGCC 541
QY 1309 AAGGTTCCGCCCTCCAGGACCCCTGAGTGTTCGCCCCACCTGTCTGCCATTTCTACGACGGCA 1368
Db 542 AAGGTTCCGCCCTCCAGGACCCCTGAGTGTTCGCCCCACCTGTCTGCCATTTCTACGACGGCA 601
QY 1369 TCTGCAAAATGGGAGGAGGCGAGTCCCAACGCCCTGTGTGCACGTGGGCTGGGCCACCTTTC 1428
Db 602 TCTGCAAAATGGGAGGAGGCGAGTCCCAACGCCCTGTGTGCACGTGGGCTGGGCCACCTTTC 661
QY 1429 TTGTGCACTCCCTAGGCGGCTTTTGAGGGACAGGTGCGGCGAGAAGGTGCCCATAGTGA 1485
Db 662 TTGTGCACTCCCTAGGCGGCTTTTGAGGGACAGGTGCGGCGAGAAGGTGCCCATAGTGA 718

RESULT 7
LOCUS AL521542/c 898 bp mRNA EST 13-FEB-2001
DEFINITION AL521542 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB003YA04 3
prime, mRNA sequence.
ACCESSION AL521542
VERSION AL521542.1 GI:12785035
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 898)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB003YA04"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
```


QY 1600 AGCCAGCACTGGACAAGGGCTGGGACCGCCGAGGTGCAGTGTTCAGGACCCCGCCGGA 1659
Db 301 AGCCAGCACTGGACAAGGGCTGGGACCGCCGAGGTGCAGTGTTCAGGACCCCGCCGGA 360
QY 1660 AGCCTCTGGGACTGTCGCTGGGCACACCGCCGAGGCTGAGGTGCGACGCGCTCAGG 1719
Db 361 AGCCTCTGGGACTGTCGCTGGGCACACCGCCGAGGCTGAGGTGCGACGCGCTCAGG 420
QY 1720 TGGCAGCACCCGAGCATCACCAACCGCCGAGGCTGAGGTGCGACGAGTGAGA 1779
Db 421 TGGCAGCACCCGAGCATCACCAACCGCCGAGGCTGAGGTGCGACGAGTGAGA 480
QY 1780 AGATGAAGGCGATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGC 1839
Db 481 AGATGAAGGCGATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGC 540
QY 1840 TGCAACTCAGGCACAGTGCAGTGCAGATGAAGAG - CAGAAAGTGTCACCCCTAGT 1898
Db 541 TGCAACTCAGGCACAGTGCAGTGCAGATGAAGAG - CAGAAAGTGTCACCCCTAGT 600
QY 1899 GACTACACTCTGTCTTCTCAAGCGGCGAGCGCAAGGCTCTGAACTACTTGGGGACTTC 1958
Db 601 GACTACACTCTGTCTTCTCAAGCGGCGAGCGCAAGGCTCTGAACTACTTGGGGACTTC 660
QY 1959 GGACCGCTTGTGGGGACCCAGGCTCGCCCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 2018
Db 661 GGACCGCTTGTGGGGA - CCAGGATCGCAGTAAAGTCCCGCAACTCTGAGCCCATGTTCTGC 719
QY 2019 CC 2020
Db 720 CC 721

RESULT 9

BF980770 966 bp mRNA EST 23-JAN-2001
LOCUS 602303938F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4395417 5',
DEFINITION mRNA sequence.
ACCESSION BF980770
VERSION BF980770.1 GI:12383582
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10092 row: 1 column: 10
High quality sequence start: 3
High quality sequence stop: 692.
Location/Qualifiers
1. .966
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4395417"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: Noti; Site_2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life

FEATURES
source

RESULT 10
BG476263
LOCUS 602525128F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4643383 5',
DEFINITION mRNA sequence.
ACCESSION BG476263
VERSION BG476263.1 GI:13408542
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 208 a 280 c 301 g 177 t
ORIGIN
Query Match 23.1%; Score 641.4; DB 172; Length 966;
Best Local Similarity 96.0%; Pred. No. 4.9e-140;
Matches 692; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

QY 349 TGCCGACCGTGTATCATCGCGCCCTCTATGCCGCTTACCGCCGAGGTGCGAGCGG 408
Db 1 TGCCGACCGTGTATCATCGCGG - CCTCTATGCCGCTTACCGCCGAGGTGCGAGCGG 59
QY 409 CCCTCGACCTGTCCCTCTATCTCCGAGAGGGGTGCTCCAGCGTGAAGTGGTGAAGA 468
Db 60 CCCTCGACCTGTCCCTCTATCTCCGAGAGGGGTGCTCCAGCGTGAAGTGGTGAAGA 119
QY 469 AGTCTCCGATGTCATATGGAACAGCGCTCAGCGCTCTACTTCAAGGATCGGCCACCA 528
Db 120 AGTCTCCGATGTCATATGGAACAGCGCTCAGCGCTCTACTTCAAGGATCGGCCACCA 179
QY 529 TCAGTCCCTCTTCAGCTTCATCACAGGACCAATTTGGACAGCTCCGGTGTGGCCTTTG 588
Db 180 TCAGTCCCTCTTCAGCTTCATCACAGGACCAATTTGGACAGCTCCGGTGTGGCCTTTG 239
QY 589 CTGTGTTGGGGCTGCGCAGGCGCTGGGTCTCCGGGATGTCACCTCGCCCTGTCTCAGG 648
Db 240 CTGTGTTGGGGCTGCGCAGGCGCTGGGTCTCCGGGATGTCACCTCGCCCTGTCTCAGG 299
QY 649 ATCATGCCCTGGGTAGTGTGTTGGGCCCAATGGGAGCAGACAGCTGAGGTCACTTGGCAGC 708
Db 300 ATCATGCCCTGGGTAGTGTGTTGGGCCCAATGGGAGCAGACAGCTGAGGTCACTTGGCAGC 359
QY 709 GCAAGGGCAACGAGGACCGCAGGGGCGCAGACAGTCAATGCCGTGTGGCTGAGCGGAGCT 768
Db 360 GCAAGGGCAACGAGGACCGCAGGGGCGCAGACAGTCAATGCCGTGTGGCTGAGCGGAGCT 419
QY 769 GCGTGTACCTGAAAGGATCATATGCGCTGTGACCCCAAGATGGAGGTGCGCTTCATGG 828
Db 420 GCGTGTACCTGAAAGGATCATATGCGCTGTGACCCCAAGATGGAGGTGCGCTTCATGG 479
QY 829 TGTGTGCCATCAACCCCTTCACCTTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGC 888
Db 480 TGTGTGCCATCAACCCCTTCACCTTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGC 539
QY 889 AGCAGAGCTGTCTGCTGCTCTATGACCTGGGACATCTGGAAGATACCCCATGCGCT 948
Db 540 AGCAGAGCTGTCTGCTGCTCTATGACCTGGGACATCTGGAAGATACCCCATGCGCT 599
QY 949 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCGCCACCCCTGGCCGCGACAGCCACTCA 1008
Db 600 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCGCCACCCCTGGCCGCGC -- AGACCACCTCA 657
QY 1009 CCCTCTACCAACAGGCGATTTGCCCTAGCCCAAGACCTACTATCGGGATGAACACATCTACC 1068
Db 658 CCTCTTACCAAGGCGATGCT ---- CAGCAAGACCTACTATCGGGATGAACACCTTTAGC 712
QY 1069 C 1069
Db 713 C 713

TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgaps-r@mail.nih.gov Tissue Procurement: ATCC/DCTD/DTp cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: http://image.llnl.gov Plate: LCLM1247 row: c column: 11 High quality sequence stop: 699.	
FEATURES		Location/Qualifiers source 1. 709 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4551874" /clone_lib="NIH_MGC_20" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected ~500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."	
BASE COUNT		151 a 214 c 253 g 91 t	
ORIGIN		Query Match 22.3%; Score 617.2; DB 153; Length 709; Best Local Similarity 97.3%; Pred. No. 2.2e-134; Matches 691; Conservative 0; Mismatches 13; Indels 6; Gaps 6	
QY	1092	CAC	TGCTCGCAACCGCAATGTGCGGAA-GCCCTCGAGGCTTGGCGACACGGCCTGTT
Db	2	CAC	TGCTCGCAACCGCAATGTGCGGAAAGGCCCTCGAGGCTTGGCGACACGGCCTGTT
QY	1151	CAT	CAGGACTACAACTACTGCGGGAAGACGAGAGATCTACAGGAGTTCCTTTGAAGT
Db	62	CAT	CAGGACTACAACTACTGCGGGAAGACGAGAGATCTACAGGAGTTCCTTTGAAGT
QY	1211	AGC	CAATGATGTCATCCCAACCTGCTGAAGGAGCGACCGAGTTCCTTTGAAGT
Db	122	AGC	CAATGATGTCATCCCAACCTGCTGAAGGAGCGACCGAGTTCCTTTGAAGT
QY	1271	GGG	CGGCGGGGAGCAAAAGCCAGGCGACCCAGAGGCTTCGCGCTTCCAGGACCC
Db	182	GGG	CGGCGGGGAGCAAAAGCCAGGCGACCCAGAGGCTTCGCGCTTCCAGGACCC
QY	1331	TGAG	TGCTTCGCGCCACCTGCTGCGATTCACGAGGATCTGCAATGGGAGGAGCGAG
Db	242	TGAG	TGCTTCGCGCCACCTGCTGCGATTCACGAGGATCTGCAATGGGAGGAGCGAG
QY	1391	TCCC	ACGCGCTGCTGTCACGCTGGGCTGGCCACCTTCTTCTGTCAGTCCCTTAGGCCGTTT
Db	302	TCCC	ACGCGCTGCTGTCACGCTGGGCTGGCCACCTTCTTCTGTCAGTCCCTTAGGCCGTTT
QY	1451	TGAG	GACAGGCTGCGCGCAGAAAGTTCGCATAGTAGGCGAGAGCGCGGCGCGGCGAG
Db	362	TGAG	GACAGGCTGCGCGCAGAAAGTTCGCATAGTAGGCGAGAGCGCGGCGCGGCGAG
QY	1511	CGAG	GAGCGCTGGGCGAGGAAGCCCGGAAGCGCGCGGCGGCGGCGGCGGCGGAGTC
Db	422	CGAG	GAGCGCTGGGCGAGGAAGCCCGGAAGCGCGCGGCGGCGGCGGCGGCGGAGTC
QY	1571	CAAG	CCAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
Db	482	CAAG	CCAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	2772	100.0	2772	5	US-09-948-941-5	Sequence 5, Appli	
2	2772	100.0	2772	5	US-09-380-337-1	Sequence 1, Appli	
3	2707.8	97.7	2767	5	US-09-948-941-135	Sequence 125, App	
4	2580.2	93.1	3534	8	US-60-278-561-1743	Sequence 1743, Ap	
5	2046.6	73.8	2168	6	US-09-758-466-258	Sequence 258, App	
6	1312.2	47.3	9180	5	US-09-380-337-3	Sequence 3, Appli	
7	1312.2	47.3	9180	7	US-09-880-107-3461	Sequence 3461, Ap	
8	1253.2	45.2	8185	6	US-09-760-475-3945	Sequence 3945, Ap	
c	1253.2	45.2	11190	5	US-09-948-941-529	Sequence 529, App	
9	1253.2	45.2	11205	5	US-09-948-941-649	Sequence 649, App	
10	529.4	19.1	628	8	US-60-253-651-22376	Sequence 22376, A	
11	488.4	17.6	639	7	US-09-864-761-22111	Sequence 22111, A	
c	473.4	17.1	519	6	US-09-878-178-1236	Sequence 1236, Ap	
12	452	16.3	464	7	US-09-864-761-5337	Sequence 5337, Ap	
c	447	16.1	3534	8	US-60-278-561-1743	Sequence 1743, Ap	
14	446	16.1	2121	1	PCT-US01-14827-5297	Sequence 5297, Ap	
c	446	16.1	2121	1	PCT-US01-14827-388	Sequence 388, App	
16	440	15.9	479	6	US-09-898-888-29321	Sequence 29321, A	
17	415.2	15.0	518	8	US-60-253-652-892	Sequence 892, App	
18	397.6	14.3	599	8	US-60-253-651-22088	Sequence 22088, A	
19	376.2	13.6	433	8	US-60-252-833-33948	Sequence 33948, A	
20	376.2	13.6	433	8	US-60-252-833-33948	Sequence 33948, A	
21	368	13.3	429	8	US-60-252-833-33900	Sequence 33900, A	
22	366	13.2	414	8	US-60-252-833-33919	Sequence 33919, A	
23	323	11.7	447	6	US-09-909-629-32588	Sequence 32588, A	
24	318	11.5	391	6	US-09-904-703-12130	Sequence 12130, A	

Db 2521 tccgccccagcctatgagcatcctgagccccccctcttccctgagcaaaactggccccg 2580
QY 2581 ATCAGAGCAGGACCTCCCTCCGACCTCTGGGAACCTCCAGAGGTCCAGAGCCCATCTCG 2640
Db 2581 atcagagcaggaacctccctccgacctctggaacctcccaagaggtccagccccactcg 2640
QY 2641 GAGCATCCGGAGGAATTCGACAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTT 2700
Db 2641 gagcatcccgaggaaatctcgagagggttagagtggtgagagagcctgctctt 2700
QY 2701 CCGTCTTTGACATAGATTTATTTTCAGTTCCTCAAGAAAGATGAATACATTTGTTAAAA 2760
Db 2701 cctgttttgacatagattatttttcagttccaagaaagatgaatacattttgttaaa 2760
QY 2761 AAAAAAAAAA 2772
Db 2761 aaaaaaaaaa 2772

RESULT 2

US-09-380-337-1
: Sequence 1, Application US/09380337
: GENERAL INFORMATION:
: APPLICANT: Chandrasekharappa, Settara C.
: Guru, Siradanahalli C.
: Manickam, Pachaiappan
: Collins, Francis S.
: Emmert-Buck, Michael R.
: Debelenko, Larisa V.
: Lubensky, Irina A.
: Liotta, Lance A.
: Agarwal, Sunita K.
: Spiegel, Allen M.
: TITLE OF INVENTION: MEN1, the Gene Associated with Multiple
: Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses
: Thereof

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,337
FILING DATE: 09-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/040,269
FILING DATE: 05-MAR-1997
APPLICATION NUMBER: WO PCT/US98/04258
FILING DATE: 04-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Lockyer, Jean M.
REGISTRATION NUMBER: 44,879
REFERENCE/DOCKET NUMBER: 015280-315100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2772 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 111..1940
OTHER INFORMATION: /product= "human menin"
FEATURE:
NAME/KEY: exon
LOCATION: 1..87
FEATURE:
NAME/KEY: exon
LOCATION: 88..555
FEATURE:
NAME/KEY: exon
LOCATION: 556..764
FEATURE:
NAME/KEY: exon
LOCATION: 765..893
FEATURE:
NAME/KEY: exon
LOCATION: 894..934
FEATURE:
NAME/KEY: exon
LOCATION: 935..1022
FEATURE:
NAME/KEY: exon
LOCATION: 1023..1159
FEATURE:
NAME/KEY: exon
LOCATION: 1160..1295
FEATURE:
NAME/KEY: exon
LOCATION: 1296..1460
FEATURE:
NAME/KEY: exon
LOCATION: 1461..2764
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-337-1
Query Match 100.0%; Score 2772; DB 5; Length 2772;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGTCCGGAGCCGCGACCTAGAGATCCAGAAAGCCACAGCGAGCGCGCCGCCGCTGA 120
Db 1 GGTGTCCGGAGCCGCGACCTAGAGATCCAGAAAGCCACAGCGAGCGCGCCGCCGCTGA 120
QY 61 ACTATTTCAGAGCTCTCCGGGCGAGGGGCGCGCCGCCGCCGCGCCGCTGAG 180
Db 61 ACTATTTCAGAGCTCTCCGGGCGAGGGGCGCGCCGCCGCCGCGCCGCTGAG 180
QY 121 AGCGCGCCCAAGAGAGCTGTTCCTCCGCTGCGCTCCATCGACGAGCTGTGCTGTTG 360
Db 121 AGCGCGCCCAAGAGAGCTGTTCCTCCGCTGCGCTCCATCGACGAGCTGTGCTGTTG 360
QY 181 CTCCCGAGCTGGGCGCGAGAGCGGACCTGGTGCTCTCTTCTTCTTGGTGCTGGGCTTCG 240
Db 181 CTCCCGAGCTGGGCGCGAGAGCGGACCTGGTGCTCTCTTCTTCTTGGTGCTGGGCTTCG 240
QY 241 TGGAGCATTTTCTGGCTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
Db 241 TGGAGCATTTTCTGGCTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
QY 301 AGCGCGCCCGCCCGCCCGAGCGCCGCTGCGGCGCTCACTACTTCCGTCGCGGACCTGT 360
Db 301 AGCGCGCCCGCCCGCCCGAGCGCCGCTGCGGCGCTCACTACTTCCGTCGCGGACCTGT 360
QY 361 CTATCATCGCGCCCTCTATGCGCGTTTACCGCGCCAGATCCGAGGCGCGCTGACCTGT 420
Db 361 CTATCATCGCGCCCTCTATGCGCGTTTACCGCGCCAGATCCGAGGCGCGCTGACCTGT 420
QY 421 CCCTCTATCTCGAGAGGGGGTGTCTCCAGCGCTGAGCTGGTCAAGAGGTCTCCGATG 480
Db 421 CCCTCTATCTCGAGAGGGGGTGTCTCCAGCGCTGAGCTGGTCAAGAGGTCTCCGATG 480


```
|||||
Db 2641 GAGCATCCGGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAGCCTGATCTCTT 2700
QY 2701 CTTGTTTTGTACATAGATTTATTTTTCAGTCTCCAGAAAGATCAATACATTTTGTAAATA 2760
Db 2701 CTTGTTTTGTACATAGATTTATTTTTCAGTCTCCAGAAAGATCAATACATTTTGTAAATA 2760
QY 2761 AAAAAAAAAA 2772
Db 2761 AAAAAAAAAA 2772
```

RESULT 3

```
US-09-948-941-125
; Sequence 125, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 2767
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-125
```

```
Query Match 97.7%; Score 2707.8; DB 5; Length 2767;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2764; Conservative 0; Mismatches 2; Indels 5; Gaps 5;
```

```
QY 1 GGTGTCGGAGCGCGGACCTAGAGATCCAGAGGCCACAGCGAGCGCGCGCGCGCC 60
Db 1 ggtgtcggagcgcgacccatagatccagagccacagcagcagcgcgcgcgcgcgc 60
QY 61 ACTATTTCCAGGCTCTGCGGGGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGTGA 120
Db 61 actatttcaggctctgcggggagggcgccgcgcgcgcgcgcgcgcgcgcgcgcgc 120
QY 121 AGCGCGCCAGAGAGAGCTCTGCCGCTGCGCTCCATCGAGAGCTGCGTGGCTGTTG 180
Db 121 agcgccgagagagcgtgtcccgctgcgctccatcgagcgcgtggtgcgcgtgtg 180
QY 181 CTGCCAGCTGGGCGGAGAGAGCGGACCTGGTGTCTCTTCTTGGTGTGCGGCTTCG 240
Db 181 ctgccagctggcgagagagagcgacacctggtgctccttcttctggtggtcgtcg 240
QY 241 TGGAGCATTTTCTGGCTGTCAACGGGTATCCCTACCAAGCTTCCGAGCTCACCTTC 300
Db 241 tggagcatcttctgctgtaacgcgctcctcctacacagttcccgagctcaccttc 300
QY 301 AGCCAGCGCGCGCGCGCGCGCGCGCTGCGGCGCTCACTACTTTCCTGCGCGGACCTGT 360
Db 301 agccagcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
QY 361 CTATCATCGCGCGCGCTATGCGCGCTTTCACCGCGCGAGATCCGAGCGCGCGTGCACCTGT 420
Db 361 ctatcatcgcgccctctatgcccgcgttcacccgcagatcccgagcgcgcgacacctg 420
QY 421 CCCTCTATCTCGAGAGGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGGTCTCCGATG 480
Db 421 ccctctatctcgagaaggggtgtctccagcgcgtgagctggtgaagaaggtctccgatg 480
QY 481 TCATATGGAACAGCCTCAGCGCGCTCTACTTCAAGGATCGGGCGCGCGCGCTCTCT 540
Db 481 tcatatggaacagcctcagcgcgtcctacttcaaggatcgggcgccacatccctct 540
```

```
QY 541 TCAGCTTCATACAGGACCAAAATGGACAGCTCCGTTGGCTTGTGCTGTGTTGGG 600
Db 541 tcagcttcacacagggaccaaattggacagctccggtgtggccttctgtgtgtgtggg 600
QY 601 CTTGCCAGCGCCTGGTCTCCGGGATGTCACCTCGCCCTGCTGAGGATCATGCTTGGG 660
Db 601 cctgcccagcctgggtctccgggagtccaacctcgccctgtctgaggtcatgctgtgg 660
QY 661 TAGTGTTTGGGCCCAATGGGGAGCAGACAGCTCAGGTTCACCTGGCAGCGCAAGGCAACG 720
Db 661 tagtgtttggggcccaatggggagcagacagctgaggttcaacctggcagcgcaaggcaacg 720
QY 721 AGGACCGCAGGGGCCAGACAGTCAATGCCGGTGTGCTGAGCGAGCTGCTGTACTGA 780
Db 721 aggaaccgaggggccaagacagtcaatgcccgtgtggctgagcgagctggtctgtacctga 780
QY 781 AAGGATCATACATGCGCTGTGACCGCAAGATGGAGTGGCGTTTCATGCTGTGCCATCA 840
Db 781 aaggatcatacatgctgtgacccgaagtggaggtggcgttcatggtgtgtgacctca 840
QY 841 ACCCTTCCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCACGACAAGCTGC 900
Db 841 acccttccattgacctgacacacgactcgtggagcttctgagctgacagaaagctgc 900
QY 901 TCTGGCTGCTCTATGACCTGGGACATCTGAAAGAGTACCCCATGGCTTAGGAAACCTGG 960
Db 901 tctggctgctctatgacctggacatctggaaaggttaccacctggccttagggaacctgg 960
QY 961 CAGATCTAGAGGAGCTGGAGCCCAACCTCGCGCGCGCAGACCCTACCTCTACCCACA 1020
Db 961 cagatctagaggctggagcccccacctggcgcgccagaccacctcaccctctaccaca 1020
QY 1021 AGGCGATTGCTCAGCAAGACCTACTATCGGATGAACACATCTACCCCTACATGTACC 1080
Db 1021 agggcattgctcagcgaagacctactatcggatgaacacatctaccctctacatgtacc 1080
QY 1081 TGCTGGCTACCACTGTGCAACCGCAATGTGCGGGAGCCCTGCGAGGCTTGGGCGGACA 1140
Db 1081 tggctggctaccactgtcgcaacgcgcaatgtcgggagacccctgacggcctggggcgaca 1140
QY 1141 CGGCCACTGTCTCCAGGACTACAATCTGCGGGAGAGGAGGAGGAGATCTACAGAGAT 1200
Db 1141 cggccactgtctccaggactacaactactgcgggagacgagcagagagatctacaaggagt 1200
QY 1201 TCTTTGAAGTAGCAATGATGTCTCCCAACCTGCTGAAAGGAGGCGAGCTTGTGTTGG 1260
Db 1201 tctttgaagttagcaatgatgtcatcccaacctgtctgaaggagcgacagcttctgtg 1260
QY 1261 AGCGGGCGAGGAGCGCGCGGGAGCAAAAGCAGGCGACCCAGAGCCAGAGTTCCGCC 1320
Db 1261 agcgggcgagagcgcgcggggagcaaaagccagggcaaccagagcgaaggttccgccc 1320
QY 1321 TCCAGGACCTGTAGTGTCTCGCCACCTGCTGCGATTCAGGAGGCTCTGCAAAATGGG 1380
Db 1321 tcagagacctgtgctgtcccaacctgtgcgacttctacagcgccatctgcaaaaggg 1380
QY 1381 AGGAGGCGAGTCCACACGCTGTGCTGACGCTGGGCGTGGGCGCACCTTCTTGTGCGAGTCC 1440
Db 1381 aggaaggcagtcaccaagcgtgtgctgcaagtggcgtggccaccttcttctgtcagtcacc 1440
QY 1441 TAGCCGCTTTTGGAGGACAGTGTGCGCAGAAAGTGTGCGCATAGTGAGCGGAGCGCGAGG 1500
Db 1441 taggcgcttttgaggacaggtgcggcagaaagtgcgcatagtgagcgcgagagcgaggg 1500
QY 1501 CGGCCGAGCGGAGGAGCGGTGGGCGGAGGAGAGCCCGGGGAAGGCGCGCGCGCGCGCCAC 1560
Db 1501 cgggcgagggcgagagcgtggggcgaggaagcccgaggagcgcgcgcgggcgccac 1560
QY 1561 GCGGGGAGTCCAAAGCAGAGGAGCGCCCGCGCCCAAGAGCCAGCAGCTGTGCAGAGGCGCC 1620
Db 1561 ggcgggagtcgaagcagagagagagcccccgccgcccgaagagcagcagctgtgacaaaggcc 1620
QY 1621 TGGCACCGCGCGCGGCTGAGTGTGAGGACCGCCCGCGGAGAGCCTCTCTGGGACTGTCTG 1680
```

Db 1621 tggcaccgcccaggggtgacgtgtcagagcccccccggaagcctcctgaggactgcgtg 1680
Qy 1681 GCAGAGCCGAGGCGCTCAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCTGAGCATCAC 1740
Db 1681 gcaagcccagggccctggaaggtgacagcagcgtcaggtgagcccccagcaccac 1740
Qy 1741 CACGCGGAGGGTCCAGTGTCTACTTCCAGAGTGAAGATGAAGGCATGAAGGAGC 1800
Db 1741 caccgcccaggggtccaggtgctcactttccagagtgagaaagatgaaagggcatgaaaggc 1800
Qy 1801 TGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGC 1860
Db 1801 tgcctgggtggccaccagaatcaactcagcagccatcaagctgcaactcaagggcacagtcgc 1860
Qy 1861 AAGTGCAGATGAAGACAGAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCA 1920
Db 1861 aagtgcagatgaagaagcagaaagtgtccacccctagtgaactacactctgtcttctctca 1920
Qy 1921 AGCGGACGCGAAAGGCTCTGAAGTACTTGGGAGCTTGGGACCCAGG 1980
Db 1921 agcggcagcgcaaaagccctcgaactactgagggacttcggaccgctgtgaggacccag 1980
Qy 1981 CTCGG-CCTTAGTCCCCAACTCTGAGCCCATGTCTGCCGCCAGCCCAAGGGGACAGG 2039
Db 1981 ctcgccccttagtcccccaactctgagcccatgtctgccccagccccaaaggagggacag 2040
Qy 2040 CCTCACCTCTACCAAAACCTTAGTCTCCGGTCCGGAGTACAGTCTGTATCAAAACCCAGC 2099
Db 2041 cctcacctctacccaacccctaggtctccgggtcccgagtagcagctgttatcaaacccag 2100
Qy 2100 ATTTTCTCAGCTCAGAACCCAGGCGCTGTGCCAGTGGTGTAGAAATAGTCTCTTCTC 2159
Db 2101 atttctccagctcagaaccagggctctgccccagctgttagaatatagttctcttc 2160
Qy 2160 CCAGAAATCCAGCGCGCAATGGAACCTCAGCGTGGGTCTTAATTACCACTCTTTAAAG 2219
Db 2161 ccagaaatccagcgggccaaatggaaacctcagctggtgtcctaattaccagctctttaa 2220
Qy 2220 GCCAGGCGCTTAGAAACCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGC 2279
Db 2221 gccagccccctagaaaccccaagctctctctcctcctcctcctcctcctcctcctc 2280
Qy 2280 TTACTAGGGTCTCTCCAGCTGTAGAGCTGAGGTTCACCCCTTAACCCAGGAGCA 2339
Db 2281 ttaactcaggctctctccagctgttagagctgaggtttcacccttaaccccaa-ggagca 2339
Qy 2340 CAGGTCCCACCTCCAGCGCGGGAGCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCG 2399
Db 2340 caggtcccacctccagccc-gggagccttaggaccactcagccccctagagtatattccg 2398
Qy 2400 CACTTCAGAAATTCATATCTTGGCAATCCAAGCTCCCTGCCCCCAAAATCACTTCAGTCTG 2459
Db 2399 cacttcagaaatccatattcttggaatccaaagctccctcctcctcctcctcctcctc 2458
Qy 2460 CTTCCAGAAATTTGGAATTCCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2519
Db 2459 c-tccagaatttgaaatccctagttctctctctctctctctctctctctctctctctct 2517
Qy 2520 CTCGCCCCAGCTATGAGCATCTTGAGCCCCCGCTCTTCTCTGACGAAACTGGCCCCG 2579
Db 2518 ctcgccccagcctatgagcatctgagcccccgccccctctctctgagcaaacctggcccc 2577
Qy 2580 GATCAGAGCAGGACTCTCCTTCCGACCTCTGGGAACCTCCAGAGGTCCAGGCCATCTC 2639
Db 2578 gatcagagcagacctccctccagccctctggaaacctccagaggtccagcccatctc 2637
Qy 2640 GGAGCATCCCGAGGAATCTGACAGAGGGGTAGAGTGGGTGACAGAGCGCTGATCTCT 2699
Db 2638 ggagcatcccgaggaaatctgcaga-gggtagagtggttgacaaagagcgtgatctct 2696
Qy 2700 TCCCTTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGAATATTTGTTAAA 2759
Db 2700 TCCCTTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGAATATTTGTTAAA

Db 2697 tctgtttttgacatagatttatttttcagttccaagaagaatgaatcacattttgttaaa 2756
Qy 2760 AAAAAA 2770
Db 2757 aaaaataaa 2767
RESULT 4
US-60-278-561-1743
: Sequence 1743, Application US/60278561
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
: FILE OF INVENTION: Polymorphisms Identified Thereby
: FILE REFERENCE: GX-0012-1 P
: CURRENT APPLICATION NUMBER: US/60/278,561
: CURRENT FILING DATE: 2001-03-23
: NUMBER OF SEQ ID NOS: 15598
: SOFTWARE: PERL Program
: SEQ ID NO 1743
: LENGTH: 3534
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No: 074428.3
US-60-278-561-1743
Query Match 93.1%; Score 2580.2; DB 8; Length 3534;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 2706; Conservative 0; Mismatches 43; Indels 10; Gaps 9;
Qy 2 GTGTCCGAGCGCGGACCTTAGATCCAGAACCCAGCAGCGCAGCGCGCGCGCGCA 61
Db 782 gcgacctagggcgggacttcatgtccacagaggtccggcgcggtgcgcgcgggtgc 841
Qy 62 CTATTTCCAGGCTGTGCGGGCAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 121
Db 842 ctagt--gtggagatgtagctcgatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 899
Qy 122 GCGCGCCAGAGAGCGTGTTCGCGCTGCGCTCCATCAGCAGCGTGTGCGCGCTGTTCG 181
Db -900 ggc 959
Qy 182 TGCGAGCTGGCGCGAGAGGAGCGGACCTGGTCTCTTTCTTTCTTTCTTTCTTTCTTT 241
Db 960 tgcgagctggcgcgagagggagcgagctgtgctccttctctgtgtgctgggcttct 1019
Qy 242 GGAGCATTTCTGCTGTCAACCGCTCATCCCTAGCAACGTTCCCGAGCTCACCTTCCA 301
Db 1020 ggaacattttctgctgtcaaccgcgtcatctccctaccacgcttcccgagctcacctcca 1079
Qy 302 GCCAGAGCG 360
Db 1080 gccagccccgc 1139
Qy 361 CTATCATCGCGCGCGCTCTATGCGCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 1140 ctatcatcgccgcctctatgcccgccttccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1199
Qy 421 CCCTCTATCTCTGAGAGAGGGGTGTCTCCAGCGGTGAGCTGTGAAGAGGTCTCCGATG 480
Db 1200 cctctctatctcggagaggggggtgtctctccagcgtgtgagctgtgtgagagaggtctccgagtg 1259
Qy 481 TCATATGGAACAGCGCTCAGCGCGCTCTACTTCAAGAGTTCGGGCGCGCGCGCGCGCGCG 540
Db 1260 tcataatggaacagcctcagcgctctacttcaaggaacgggccccacatccagtcctct 1319
Qy 541 TCAGCTTCATCACAGGACCAAAATTTGGACAGCTCCGGTGTGGCGCTTTTGTGTGGG 600

|||||
Db 1320 tcaagttcatacaggacacaaatlgyacagctccggtgtggtcccttgtgtggtggg 1379
QY 601 CTTGCCAGGCCCTGGTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGG 660
Db 1380 cctgcgaagccctgggtctccggatgttccacctcgccctgtcctgagatcatgctggg 1439
QY 661 TAGTGTGTTGGCCCCAATGGGGAGCAGACAGCTGAGGTACCTTGGCAGCGCAAGGCAACG 720
Db 1440 tagtgtttgggccccaatgggagcagacgtgaggtcaccttggcagcgcaagggcaacg 1499
QY 721 AGGACCGAGGGCCAGACAGTCAATGCGGTGTGGCTGAGCGGAGCT--GGTGTACCTG 779
Db 1500 aggaacgcaggggccagacagtcacatgcgggtgtgctgagcgagctgggtgtacctg 1559
QY 780 AAAGGATCATACATCGCTGTGACCGCAAGATGGAGGTGGCGTTCATGTGTGTGCCATC 839
Db 1560 aaagatcatacatgctgtgtgaccgcaaga tggaggttccatgggtgtggtgcca tc 1619
QY 840 AACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGACGTGAGCAGAGCTG 899
Db 1620 aaccttcattgacctgcacacgactcgtgtgagcttctgacctgcagcagaagctg 1679
QY 900 CTCTGGCTGTCTATGACCTGGGACATCTGGAAGGTACCCCATGCCCTTAGGGAACTGG 959
Db 1680 ctctggctgtctatgacctgggacatctggaaggtacccatggccttagggaaacctg 1739
QY 960 GCAGATCTAGAGAGCTGGAGCCACCCCTGCGCGGCAGACCCACTCACCTCTACACAC 1019
Db 1740 gcagatctagagagctggagccccccctggccggcagagccccactaccctaccac 1799
QY 1020 AAGGCATTGCTCAGCCAAAGACTACTATCGGGATGAACACATCTACCCCTACATGTAC 1079
Db 1800 aaggcattgtctcagcaagacctactatcggatgaacacatctacccttaactgtac 1859
QY 1080 CTGGCTGGCTACACTGTGCGCAACCGCAATGTGCGGGAAGCCCTGAGGCCCTGGCGGAC 1139
Db 1860 ctggctggctaccactgtgcgaacgcgaatgtgcgggaagccctgcaggcctgggcggac 1919
QY 1140 ACGGCCACTGTCTCAGGACTACAACTACTCGCGGAGACGAGAGATCTACAGGAG 1199
Db 1920 acggccactgtctatccaggactacaactactcggggagagagagatctacaaggag 1979
QY 1200 TTTCTTTGAAGTAGCCCAATGATGTCTCCCAACCTGCTGAAGGAGCAGCCAGCTGTCTG 1259
Db 1980 ttctttgaagtagccaatgattcatcccaacctgtgaaggaggaaccaagcttgtg 2039
QY 1260 GAGGCGGCGAGAGGCGCGGGGAGCAAAAGCCAGGCAACCCAGAGCCCAAGGTTCCCGCC 1319
Db 2040 gaggcgggcgaggagcgcggggagcaaaagccaggggcacccagagccaaaggttccgpc 2099
QY 1320 CTCCAGGACCTTGAGTGTTCGCCCCACCTGCTGCGATTTCTAGACGGCATCTGCAANTGG 1379
Db 2100 ctccaggacctgtgtgtctccacctgtgtgattctagaaggatctgcaaatgg 2159
QY 1380 GAGGAGGCGAGTCCACGCTGTGTGCACTGGGTGGCGCCACCTTTCTTGTGAGTCC 1439
Db 2160 gaggaggcagttcccaacctgtgtgcaagtgggtgtggccacctttcttgtgaagtcc 2219
QY 1440 CTAGGCGGTTTGGAGGACAGGTGGCGGAGAGAGGTGCGCATAGTAGGCGAGAGCGCGAG 1499
Db 2220 ctaggcggttttgaggacaggtgcgcgagaggttgcgcatagtgagcgagagcgcgag 2279
QY 1500 GGGCGGCGGCGAGGAGCGGTGGGCGAGGAAGCCCGGGAAGGCGCGGCGGGGCGCCA 1559
Db 2280 gggcgcgaggcgagagcggtggggcgaggaagcccggaagggcgggcgggggccca 2339
QY 1560 CGCGGGAGCTCCAAAGCAGAGGAGCGCCCGCCGCCCCAAGAACCACTGACAGAGGCG 1619
Db 2340 cggcggaagtccaaagccagagagagcccccgccgccccaaagaagccagacctggacaaggcg 2399
QY 1620 CTGGGCAACCGGCGAGGCTGAGTGTGAGA--CCCCCGCGGAAGCCCTCTGGGAGCTGTCCG 1678
|||||

Db 2400 ctggggaccggccaggggtgagtgtagtgacagccccccccggagagcctctctgggtatgcg 2459
QY 1679 TGGCAGAGCCCGAGGCGCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGACAGATC 1738
Db 2460 tggcacagccccgagggccctgaaagg tggcagcacggctcagg tgcceagccccgcagca tc 2519
QY 1739 ACCACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGAATGAAGGCGATGAAGGA 1798
Db 2520 accacgcgcggaggggtccagtgctcaacttccagagtgagaagatgaaggcgcatgaagga 2579
QY 1799 GCTGCTGGTGGGCCCAACAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCACAGTC 1858
Db 2580 gctgtgtgtggccaccaagatcaactcgagcgccatcaagctgcaactcacggcacagtc 2639
QY 1859 GCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTGACTACACTCTGTCTTTCT 1918
Db 2640 gcaagtgcagatgaagaagcagaaggtgtccaccctctagtgtactacactctgtcttctct 2699
QY 1919 CAAGCGGCAGCGCAAGGCTCTGAACACTACTGGGACCTTTCGGACCGCTTGTGGGGACCCA 1978
Db 2700 caagcggcagcgcaagggcctctgaactactggggaacttcggaccgcttgtggggaccca 2759
QY 1979 GSGTCCG--CCTTAGTCCCCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGAGCA 2037
Db 2760 ggctcgcccttagtcccccactctgagcccatgtcttgccccccagcccaaggagca 2819
QY 2038 GGCCTCACCTCTACCCAAACCCCTAGGTTCCCGTCCCGAGTACAGTCTGTATCAAAACCCA 2097
Db 2820 ggctcacctctacccaaaccttaggttccccggctcccgagtagagctgtgtatacaaccca 2879
QY 2098 CGATTTCCTCCAGCTCAGAAACCCAGGGCTCTGCCCGAGTCTGTAGTAATATAGTCTCTTTC 2157
Db 2880 cgattttctccagctcagaaccagggtctgtccccagtcgttagaatalaggtctcttc 2939
QY 2158 TCCAGATTCCTCAGCGCGCCCAATGGAAACCTCAGCGTGGGTCTTAATTACCAGTCTTTAA 2217
Db 2940 tcccaagatcccgagcgccaatggaaacctcacgtgggtcctaataattaccagctcttaa 2999
QY 2218 AGGCCAGCCCTTAGAAACCCCAAGCTCTCTCGGAAACCGTCCACCTAGAGCCAGACCCA 2277
Db 3000 agggccagccccctagaaaccccaagctctctctctcggaaccgctcaacctagagccagacca 3059
QY 2278 CGTTACTAGGGCTCTCTCCAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCCAAGGAG 2337
Db 3060 cgttactcaggtctctccagctttagagagtgaggtttcaccccttaacccaa--ggag 3118
QY 2338 CACAGTCCCACTCCAGCGCGGGAGCCTAGGACCACTCAGCCCTAGAGTATATTTC 2397
Db 3119 cacaggtcccaactccagccc--ggagcctaggaaccactcagccccctagagtatatttc 3177
QY 2398 CGCACTTCAGAAATCCCATATCTTGCGAATCCCAAGCTCCCTGCCCAAAATAACTTCAGTCT 2457
Db 3178 cgcactcgaatctccatctcttgcaatcccaagctccccctgcacccaaataactctcagtc 3237
QY 2458 TGTCTCCAGAAATTTGAAATCTTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2517
Db 3238 tgc-tccagaatttgaaatctctgttctctctctctgtatccccgagctctgggacaaa 3296
QY 2518 ACTCGGCGCCCGAGCCCTATGAGCATCTCAGCCCGCCCTCTTCTCTGACGAACCTGGGCC 2577
Db 3297 aactcgcccccgagcctatgagcatctgagccccctctctctctctctctctctctctctct 3356
QY 2578 CGGATCAGAGCAGGACCTCCCTTCCGACCTCTGGAACCTCCACAGGTGCCAGCCCATC 2637
Db 3357 cggatcagagcgaggacctctctccacctctcggaacctctcggaacctccagaggtccagccatc 3416
QY 2638 TGGAGCATCCCGGAGGAAATCTGCAAGAGGGTGTAGAGTGGGTGACAAAGAGCTGATCT 2697
Db 3417 tcggagcatcccgaggaatctgcaga--gggttagagtggtgacaaagagcctgatct 3475
QY 2698 CTTCTCTGTTTGTACATAGATTATTTTTCAGTTCACGAAGAAATGAATACATTTGTT 2756
Db 3476 ctctcgtttgtacatagattttttcagttcccaagaaagatgaatacatatttctgt 3534

RESULT 5
US-09-758-466-258
; Sequence 258, Application US/09758466
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM036
; CURRENT APPLICATION NUMBER: US/09/758,466
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 814
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
; LENGTH: 2168
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-758-466-258

Query Match 73.8%; Score 2046.6; DB 6; Length 2168;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2112; Conservative 3; Mismatches 2; Indels 6; Gaps 6;

QY	651	CATGCCCTGGGTAGTGTGGGCCCCAATGGGAGCAGACAGCTGAGGTCACTGGCAGGCG	710		691	ggttccgcctccaggaacctgagttctcgccacacctgctgagattctacgatlgygcac	750
Db	32	cagggctggtagtgtttg9ggccaa1tgg9gagcagacagtgaagtcacctggcacggc	91		1371	TGCAATGGGAGGAGGCGCAGTCCACACCCCTGTGCTGCACCTGGGCTGGCGCCACTTTC	1430
QY	711	AAGGCCAACGAGGACCCGAGGGCCACACAGTCAATCCCGTGTGGCTGAGCGGAGCTGG	770		751	tgcaaatgggagggagggcagtcctccacgctgtgctgcagtgggctgggacaccttctt	810
Db	92	aagggcaacagagaccgaggggccaacagtcactcaatccggtgtgctgagcgagctg	151		1431	GTCCAGTCCCTAGGCGCTTTTGGAGGACAGGTGCGGAGAGGTGCGCATAGTTCAGCCGA	1490
QY	771	CTGTACTGAAAGGATCATACATCGCTGTGACCGCAAGATGAGGTGGGTTCATGGTG	830		811	gtcagtcctcctagggccgttttgagggacaggtgcg9cagaaggtgcacatag1gagccga	870
Db	152	ctgtacgtgaaagatcatatcatgctgtgacccgaagatggaggtggcggttcatggt	211		1491	GAGCCGAGGCGGCGGAGCGGAGGAGCCGTGGGGGAGGAGCCCGGAGGAGCCGGCGG	1550
QY	831	TGTGCCATCAACCTTCCATTGACCTGCACACGACGCTGCTGGAGCTTCGACGCTCAG	890		871	gagggcagagcg9ccgagggccgtg9ggcgagagcccg9gagcccg9gagccg9c9g	930
Db	212	tg9ccatcaaccttccattgacctgtcacccgacctgcgtggagcttctg9cagctgag	271		1551	CGGGGCCCGGCGGAGTCCAAGCCAGAGAGCCCGCCCGCCCGCCCAAGAGCCAGCACTG	1610
QY	891	CAGAAGTGTCTGTGGCTCTATGACCTGGGACATCTGGAAGAGGTACCCCATGGCCTTA	950		931	c9ggggccacg9cg9gagtcacagccagagggagcccccgc9gcccagaagagcagcactg	990
Db	272	cagaagctgtctggtgctctatgacctggacatctggaaggttaccctcagggccta	331		1611	GACAAGGGCGCTGGGCACCGGCGCAGGGTGCAGTGTACAGACCCCGCCCGGAAAGCTTCT	1670
QY	951	GGGAACCTGGCAGATCTAGAGGAGCTGGAGCCACCCCTGGCGGCGCAGACCACTCACC	1010		991	gacaagggcctgg9ccaccg9ccaggtgcaggtgtcagagacccccccg9aagcctcctg9g	1050
Db	332	gg9aacctggcagatctagaggagctggagccacccctggccg9ccagacccactcacc	391		1671	ACTGTGCTGGCAGACCGCCGAGGCCCTGGAAGGTGGCAGCAGCGCTCAGGTGCAGACACC	1730
QY	1011	CTCTACCAAGGCAATGTGCTCAGCCCAAGACCTACTATCGGATGAACATCTACCCC	1070		1051	actgtcgtggcacagcccgagggccctgaaggtggcagcag9gctcaggtgcagacccc	1110
Db	392	ctctaccacaagggcattg -ctcagccaagacctactatcctgg9gatgaacacctacccc	450		1731	GCAGCATCACACCGCGGAGGTCCAGTGTCTCAGTTCCTCAGAGTGAGAGATGAAGSGC	1790
QY	1071	TACATGTACCTGGCTGCTACCTGCTCGCAACCGCAATGTGGGGGAAGCCCTGCAGGCC	1130		1111	gc9gcatcaacccacg9cg9gggttccaggtctcactctccag9gtg9aag9g9a	1170
Db	451	tacatgtacctggctggttaccactgtcgcaaccgcaatgtg9gggagccctgcagggc	510		1791	ATGAAGGAGTGTGTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGC	1850
QY	1131	TGGCGGACAGCGGCATCTATCCAGGACTACACACTACTGCTGCGGGGAAGCAGGAGATC	1190		1171	atgaagagctgtggtggccaccaagatcaactcgagcgccatcaagctgcgaactcacg	1230
Db	511	tgg9cgacacggccactgtcatccaggactacaactactgcccggggaagcagggagatc	570		1851	GCACAGTCGCAAGTGCAGATGAAGACGACAAAGTGTCCACCCCTAGTACTACACTCTG	1910
QY	1191	TACAAGGAGTCTTTGAAGTAGCCCAATGATGTATCCCAACCTGCTGAAGGAGGAGCC	1250		1231	gcacagtgc9caagtgcagatg9aagagc9aag9gtcccacccctagtgactcacactcg	1290
Db	571	tacaagagttcttctgaag1agccaa1ga1gctcatccccaacctgtcga9gg9gagcc	630		1911	TCCTTCTCAAGCGGAGCGCAAGGCTCTGAACACTACTGGGACTTCGGACCGCTGTGTG	1970
QY	1251	AGCTTGTGGAGGGCGGAGGAGCGCGGGGAGCAAAAGCCAGGCGACCCAGCCAGCCAA	1310		1291	tcttctcctcaagc9g9cagc9maaaagcctcgaactactg9g9gactc9g9ac9c9ctg	1350
Db	631	agctgtgrag9cg9cgagagc9cg9gggagcaagccag9gcaagccacccagagccaa	690		1971	GGGACCCAGGCTCCG - CTTTAGTCCCCCAACTCTGAGCGCCATGTTTCGCCCCAGCCAA	2029
QY	1311	GGTTCGCCCTCCAGGACCCCTGAGTGTCTCGCCCACTGTGCGGATCTAGGAGGCGATC	1370		1351	gg9acccag9ctccgccccttagtcccccaactctgagcccatgtctctgccccagcccaa	1410

Db 1769 tatatttcgcacttcagaattccatattcttgcgaatcccaagctccctcgtcccccataaac 1828
QY 2450 TTCAGTCCTGTCCTCCAGAAATTTGAAATCCTAGTTCTCTCTCTCGTATCCCGAGTCTG 2509
Db 1829 ttccagtcctgc-tccagaatttggaatcctcagttctcctctcctcgtatcccgagtcg 1887
QY 2510 GCACACAAACTCCCGCCCGCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAA 2569
Db 1888 ggacacaaactccgccccagcctatgagcatcctgagcccgccctctcctgacgaa 1947
QY 2570 ACTGGCCCGGATCAGAGCAGCAGCCTCCCTCCGACCCCTCTGGAACTCCCGAGAGTCC 2629
Db 1948 actgcccccgatcagcagcagcctccctccgacctctgggaacctcccgagggctcc 2007
QY 2630 AGCCCATCTCGAGCATCCCGGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAG 2689
Db 2008 agcccatctcgagcatcccgagggaatctgcaga-gggttaggagtggtgacaaagag 2066
QY 2690 CTTGATCTCTCTCTGTTTGTACATAGATTTATTTTCAGTTCACAGAAAGATGAATACA 2749
Db 2067 cctgatctctcctctgtttgtacatagattttttcagttcccaagaagaatgataca 2126
QY 2750 TTTTGTAAAAAATAAAAAA 2772
Db 2127 tttgttaaaaaaataaaaaa 2149

RESULT 6

US-09-380-337-3
; Sequence 3, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Setara C.
; Guru, Siradanahalli C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Debelenko, Larisa V.
; Lubensky, Irina A.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated with Multiple
; Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses
; Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,337
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-315100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..9180
; OTHER INFORMATION: /note= "genomic sequence for MEN1 gene"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1680..1766
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1767..2264
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2265..2732
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2733..4296
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4297..4505
; FEATURE:
; NAME/KEY: intron
; LOCATION: 4506..4715
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4716..4844
; FEATURE:
; NAME/KEY: intron
; LOCATION: 4845..5176
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5177..5217
; FEATURE:
; NAME/KEY: intron
; LOCATION: 5218..5297
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5298..5385
; FEATURE:
; NAME/KEY: intron
; LOCATION: 5386..6024
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6025..6161
; FEATURE:
; NAME/KEY: intron
; LOCATION: 6162..6622
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6623..6758
; FEATURE:
; NAME/KEY: intron
; LOCATION: 6759..7195
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7196..7360
; FEATURE:
; NAME/KEY: intron
; LOCATION: 7361..7577
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7578..8881
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-380-337-3

Query Match

47.3%; Score 1312.2; DB 5; Length 9180;

Best Local Similarity 96.5%; Pred. No. 1.4e-245;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1382 GGAGGCGAGTCCACGACCTGCTGCTGACGCTGGGCTGGCCACCTTTCTTCTGAGTCCCT 1441
Db 7499 GGAATGGCGAGCAGAGGCTCTGGAGTTCCAGCCACTGCGCGGAACCTTCTCTCTCACTT 7558
QY 1442 AGGCCGCTTTGAGGCGACAGGTGGCGCAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 1501
Db 7559 GCTCTCCCACTGGCCAGGTGCGCGCAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 7618
QY 1502 GGCGAGGCGAGAGCGGTGGGCGAGGAGGAGGCGGGAAGGCGGCGGCGGCGGCGGCGGCGG 1561
Db 7619 GGCGAGGCGAGAGCGGTGGGCGAGGAGGAGGCGGGAAGGCGGCGGCGGCGGCGGCGGCGG 7678
QY 1562 GCGGAGGTCCAAGCCAGAGGAGCCCGCGCGCCCAAGAACGACGACTGACAAAGGCGCT 1621
Db 7679 GCGGAGGTCCAAGCCAGAGGAGCCCGCGCGCCCAAGAACGACGACTGACAAAGGCGCT 7738
QY 1622 GGGCAGCGGCGAGGCTGACGTGTGAGGACCCCGCGGAGGCGCTTGGGACTGTGCGTGG 1681
Db 7739 GGGCAGCGGCGAGGCTGACGTGTGAGGACCCCGCGGAGGCGCTTGGGACTGTGCGTGG 7798
QY 1682 CACAGCGGAGGCGCTGAAGGTGGCAGCAGCGCTCAGGCTCCAGCAGCCCGCAGCATACC 1741
Db 7799 CACAGCGGAGGCGCTGAAGGTGGCAGCAGCGCTCAGGCTCCAGCAGCCCGCAGCATACC 7858
QY 1742 ACCCGCGAGGCTCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCT 1801
Db 7859 ACCCGCGAGGCTCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCT 7918
QY 1802 GCTGGTGGCCACCAAGATCAACTGAGCGCCATCAAGCTGCAACTCAGCGACAGTCGCA 1861
Db 7919 GCTGGTGGCCACCAAGATCAACTGAGCGCCATCAAGCTGCAACTCAGCGACAGTCGCA 7978
QY 1862 AGTGCATGAGNAGCAGAAAGTGTCCACCCCTAGTACACTGTCTTCTCTCAA 1921
Db 7979 AGTGCATGAGNAGCAGAAAGTGTCCACCCCTAGTACACTGTCTTCTCTCAA 8038
QY 1922 GCGCAGCGCAAGGCGCTCTGAAGTCTGGGACTTCGGGCGCTTGGGGACCCAGGC 1981
Db 8039 GCGCAGCGCAAGGCGCTCTGAAGTCTGGGACTTCGGGCGCTTGGGGACCCAGGC 8098
QY 1982 TCGCGCTTAGTCCCGCAACTCTGAGCCCATGTCTTGCCTCCCGAGCCCAAGGGGACAGGC 2041
Db 8099 TCGCGCTTAGTCCCGCAACTCTGAGCCCATGTCTTGCCTCCCGAGCCCAAGGGGACAGGC 8158
QY 2042 TCAGCTTACCCCAACCTAGGTTCCCGGCTCCGAGTACAGTCTGTATCAACCCAGCAT 2101
Db 8159 TCAGCTTACCCCAACCTAGGTTCCCGGCTCCGAGTACAGTCTGTATCAACCCAGCAT 8218
QY 2102 TTTCTCCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTAGATATAGTCTCTTCTCC 2161
Db 8219 TTTCTCCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTAGATATAGTCTCTTCTCC 8278
QY 2162 AGAATCCAGCGCGCCCAATGAAACCTCAGCGTGGGCTCTAATTACAGTCTTTAAAGC 2221
Db 8279 AGAATCCAGCGCGCCCAATGAAACCTCAGCGTGGGCTCTAATTACAGTCTTTAAAGC 8338
QY 2222 CCAGCCCTTAGAATCCCAAGCTCTCTCGGACCCGCTCAGCTAGAGCCAGACCAAGCTT 2281
Db 8339 CCAGCCCTTAGAATCCCAAGCTCTCTCGGACCCGCTCAGCTAGAGCCAGACCAAGCTT 8398
QY 2282 ACTCAGGCTCTCCAGCTCTGTAGGAGCTGAGGTTTACCTCTTAACCCAAAGGAGCACA 2341
Db 8399 ACTCAGGCTCTCCAGCTCTGTAGGAGCTGAGGTTTACCTCTTAACCCAAAGGAGCACA 8458
QY 2342 GGTCCCACTCCAGCCCGGGAGCGCTTAGGACCACTAGACCCCTTAGGAGTATATTTCCGCA 2401
Db 8459 GGTCCCACTCCAGCCCGGGAGCGCTTAGGACCACTAGACCCCTTAGGAGTATATTTCCGCA 8518
QY 2402 CTTCAAGATTCATATCTTGGGAATCCAAAGCTCCCTCGCCCAATAACTTTCAGTCTCTGCT 2461
|||||

Db 8519 CTTCAGAAATCCCATATCTTTCGAATCCAAAGTCCCTGCCCAAAATTAATTCAGTCTCTGCT 8578
QY 2462 TCCAGAAATTTGGAATCTCTAGTTTCTCTCTCTCGATCCGAGTCTGGGACACAAAAC 2521
Db 8579 TCCAGAAATTTGGAATCTCTAGTTTCTCTCTCTCGATCCGAGTCTGGGACACAAAAC 8638
QY 2522 CCGCCCGCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAAACTGGCCCCGGA 2581
Db 8639 CCGCCCGCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAAACTGGCCCCGGA 8698
QY 2582 TCAGAGCAGACCTCCCTCCGACCCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGG 2641
Db 8699 TCAGAGCAGACCTCCCTCCGACCCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGG 8758
QY 2642 AGCATCCCGAGGAAATCTCAGAGGCGTTAGAGTGGTGACAAGAGCTGATCTCTTC 2701
Db 8759 AGCATCCCGAGGAAATCTCAGAGGCGTTAGAGTGGTGACAAGAGCTGATCTCTTC 8818
QY 2702 CTGTTTGTACATAGATTTATTTTTCAGTTCCAAAGAAAGATGAATATTTTGTAAAAA 2761
Db 8819 CTGTTTGTACATAGATTTATTTTTCAGTTCCAAAGAAAGATGAATATTTTGTAAAAA 8878
QY 2762 AAAAAAAA 2770
Db 8879 AATATAAA 8887

RESULT 7
US-09-880-107-3461
; Sequence 3461, Application US/09880107
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3461
; LENGTH: 9180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. U93237
; NAME/KEY: unsure
; LOCATION: (1), (9180)
; OTHER INFORMATION: n - a o r c o r g o r t
US-09-880-107-3461

Query Match 47.3%; Score 1312.2; DB 7; Length 9180;
Best Local Similarity 96.5%; Pred. No. 1.4e-245;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1382 GGAGGCGAGTCCACGACCTGCTGCTGACGCTGGGCTGGGCCACCTTTCTTCTGAGTCCCT 1441
Db 7499 ggaatggcgagcagaggtctggagttccagccactgcgcggaaccttctctctcaactt 7558
QY 1442 AGGCCGCTTTGAGGCGACAGGTGGCGCAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 1501
Db 7559 gctctcccaactgccccaggtgccccaggtgccccaggtgccccaggtgccccaggtgccccaggt 7618
QY 1502 GGCGAGGCGAGGAGCGGTGGGCGAGGAGGAGGCGGGAAGGCGGCGGCGGCGGCGGCGGCGG 1561
Db 7619 ggcgagggcgagagagcgcttggggcgaggaagccccggggagggcgcgcgggggcgagggcg 7678
|||||

QY 1562 GGGGAGTCCAAAGCCAGAGAGCCCGCCGCCCAAGAACCCAGCAGCTGACAAAGGCGCT 1621
Db 7679 gggggagtgccaagcagagagagcccgccgcccgaagagccagcactgagcaaggcct 7738
QY 1622 GGGCAGCGCCAGGGTGAGTGTGAGGAGCCCGCCGGAAGCCTCTGAGACTGTCTGCTGG 1681
Db 7739 gggcagcgccaggggtgcagtgccaggaccccccccggaagcctctggactgtcgctgg 7798
QY 1682 CACAGCCCGAGGCGCTGAAGGTGGCAGCACGGCTCAGGTGCCAGCACCCCGCAGCATCAC 1741
Db 7799 cacagcccgagggcctgaagggtggcagcaagcgcctcaggtgcagcagcccgagcatcac 7858
QY 1742 ACGCCGGAGGGTCCAGTCTCTCACTTTCAGAGTGAGAGATGAAGGGCATGAAGGAGCT 1801
Db 7859 accgcccggaggggtccagtgctcacttccagagtgagagatgaaggcatgaaggagct 7918
QY 1802 GCTGGTGGCCACCAAGATCAACTCGAGCGCCCATCAAGCTGCAACTCACGGCAGCTCGCA 1861
Db 7919 gctgggtggccaccaagatcaactcgagcgccatcaagctgcaactcacgacagtcgca 7978
QY 1862 AGTGCAAGTGAAGACGACAAAGTGCCACCCCTAGTGTACTACACTCTCTTTCTCTCAA 1921
Db 7979 agtgcaagtgaagaagacagaagtgccaccctcagtgactcaactctgtcttctctcaa 8038
QY 1922 GGGCAGCGCAAGGGCTCTGAACACTACTGGGACTTCGGACCCCTTGTGGGACCCAGGC 1981
Db 8039 gggcagcgcaaggcctctgaactactgggacttcggaccgcttggggacccagcg 8098
QY 1982 TCGGCTTAGTCCCCCACTCTGAGCCCATGTCTGCCCGCCAGCCCAAGGGGACAGGCC 2041
Db 8099 tcggccttagtcccccaactctgagcccatgttctgccccagcccaaggggagcagggc 8158
QY 2042 TCACCTCTACCCAAACCCTAGTTCGCGTCCGAGTACAGTGTGTATCAAAACCCACGAT 2101
Db 8159 tcactcttaccaaacccctaggttcccggtcccgagtagcagtcgtatcaaaacccagat 8218
QY 2102 TTCTCCAGCTCAGAACCCAGGGCTGTGCCAGTGGTGTAGATATAGTGTCTTCTCC 2161
Db 8219 ttctccagctcagaacccagggctgtgccagtcggttagaataagtgctctctctccc 8278
QY 2162 AGAATCCAGCGGCCAATGGAACCTACGCTGGTGTCTTAATACCACTTTAAAGC 2221
Db 8279 agaatccccgcccgaatggaaacctcacgtgggtccctaatatccagtcctttaaaggc 8338
QY 2222 CCAGCCCTTAGAAACCCAAAGCTCTCTCTCGAAACCCCTCACCTAGACGACAGCAACGTT 2281
Db 8339 ccagcccttagaaacccaaagctctctctcgaaacgctcaactagagccagaccacgtt 8398
QY 2282 ACTCAGGGCTCTCCAGCTTGTAGAGCTGAGGTTTCCACCTTAACCCAAAGGAGCACA 2341
Db 8399 actcagggtcctcccagctttagagctgaggtttccaccttaacccaaaggagcaca 8458
QY 2342 GTTCCACCTCCAGCCCGGGAGCCTAGGACCACTCAGCCCTTAGAGTATATTCCGCA 2401
Db 8459 gttccacctccagcccgggagctcaggaacctcagcccccttaggagtagatatctccga 8518
QY 2402 CTTAGAATTCCATATCTTGGAAATCCAAAGCTCCCTGCCCCCAATTAACCTTCAGTCTGCT 2461
Db 8519 cttcagaattccatattcttggaaatccaaagctccctgccccaaataaacttcagtcctgt 8578
QY 2462 TCCAGAAATTTGAAATCTTAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2521
Db 8579 tcagaaatttgaaaatccctagttctctctctctctctctctctctctctctctctctct 8638
QY 2522 CGCCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 2581
Db 8639 cgccccccagcctatgagcatctgagcccccgcctctctctgagaaactggccccgga 8698
QY 2582 TCAGACGAGACCT 2641
Db 8699 tcagacgagacacctctctccgacctctctggaaacctccagaggtccagcccatctgg 8758
QY 2642 AGCATCCCGAGGAAATCTGCAGAGGGGTTAGAGTGGGTGACAAAGAGCCTGATCTTTC 2701

Db 8759 agcatcccgaggaataatctgcagaggggttagagtggtgacaagagcctgactctlc 8818
QY 2702 CTGTTTGTACATAGATTTATTTTTCAGTTCGAAGAAAGATGAATACATTTGTTAAAAA 2761
Db 8819 ctggtttgtacatagattatttttccagttccaagaagatgaatacatlltgitaaaaa 8878
QY 2762 AAAAAAAA 2770
Db 8879 aaataataa 8887

RESULT 8
US-09-760-475-3945/c
; Sequence 3945, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT49
; CURRENT APPLICATION NUMBER: US/09760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3945
; LENGTH: 8185
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-760-475-3945

Query Match 45.2%; Score 1253.2; DB 6; Length 8185;
Best Local Similarity 96.2%; Pred. No. 4 1e-234;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;

QY 1382 GGAGGGCACTCCACGCCCTGTGTCACGTGGGCTGGGCGACCTTTCTGTGTCAGTCCCT 1441
Db 1386 GGATGCCAGAGCAGCGGTCTTGGAGTTCAGCCACTTGGCGGCAACCTTGTCTCACTT 1327
QY 1442 AGGCCGTTTGGAGGACAGTGGCGCAGAGGTGCCATAGTCAGCCGAGAGCCGAGCCGAGGC 1501
Db 1326 GCTCTCCCACTTGGCCAGGTGGCGCAGAGGTGCCATAGTCAGCCGAGAGCCGAGGC 1267
QY 1502 GGCGAGGGCCGAGGAGCCGTGGGGCGAGGAAGCCGGGAGGCGGGCGGCGCCACG 1561
Db 1266 GGCGAGGGCCGAGGAGCCGTGGGGCGAGGAAGCCGGGAGGCGGGCGGCGCCACG 1207
QY 1562 GCGGAGTCCAGCCAGAGAGGCCCGCGGCCCAAGAGCCAGCACTGGACAGGGCCT 1621
Db 1206 GCGGAGTCCAGCCAGAGAGGCCCGCGGCCCAAGAGCCAGCACTGGACAGGGCCT 1147
QY 1622 GGGCACCGCCAGGGTGCAGTGTCCAGGACCCCGCCGAAAGCCCTCTGGGACTCTCGCTGG 1681
Db 1146 GGGCACCGCCAGGGTGCAGTGTCCAGGACCCCGCCGAAAGCCCTCTGGGACTCTCGCTGG 1087
QY 1682 CACAGCCCGAGGCGCTTAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 1741
Db 1086 CACAGCCCGAGGCGCTTAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 1027
QY 1742 ACGCCGGAGGGTCCAGTGTCTACTTTCAGAGTGAAGATGAAGGCGATGAAGGAGCT 1801
Db 1026 ACGCCGGAGGGTCCAGTGTCTACTTTCAGAGTGAAGATGAAGGCGATGAAGGAGCT 967
QY 1802 GCTGTGGCCCAACAGATCAACTCAGGCGCCATCAAGCTGCAACTCACGCGACAGTCGCA 1861
Db 966 GCTGTGGCCCAACAGATCAACTCAGGCGCCATCAAGCTGCAACTCACGCGACAGTCGCA 907
QY 1862 AGTGAGATGAAGAACAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTCTCTCAA 1921
Db 906 AGTGAGATGAAGAACAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTCTCTCAA 847
QY 1922 GCGGCGAGCGCAAGAGCGCTCTGAACACTTCTGGGAGCTTTCGAGCCGCTTGTGGGACCCAGCG 1981

```
Db 846 GCGCAGCGCAAGAGCCCTCTGAACACTCTGGGACTTCGGACCGCTTGTGGGACCCACAGC 787
Qy 1982 TCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCGCCCAAGGGACAGGC 2040
Db 786 TCCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCGCCCAAGGGACAGGC 727
Qy 2041 CTCACCTCTACCCAAACCCCTAGGTTCCGGTCCGGGTACAGTCTGTATCAAAACCCACGA 2100
Db 726 CTCACCTCTACCCAAACCCCTAGGTTCCGGTCCGGGTACAGTCTGTATCAAAACCCACGA 667
Qy 2101 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCAAGTCTGTAGATATAGTCTCTCTCC 2160
Db 666 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCAAGTCTGTAGATATAGTCTCTCTCC 607
Qy 2161 CAGAAATCCAGCGCGGCAATGGAACCTCAGCTGGGTCTTAATTACAGCTTTTAAAGG 2220
Db 606 CAGAAATCCAGCGCGGCAATGGAACCTCAGCTGGGTCTTAATTACAGCTTTTAAAGG 547
Qy 2221 CCCAGCGCTTAGAAACCCAGCTCCTCTCGGAAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 546 CCCAGCGCTTAGAAACCCAGCTCCTCTCGGAAACCGCTCACCTAGAGCCAGACCAACGT 487
Qy 2281 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGGTTTCAACCTTAACCCAAAGGAGCAC 2340
Db 486 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGGTTTCAACCTTAACCCAAAGGAGCAC 428
Qy 2341 AGGTCCACCTTCCAGCGCGGAGCCTAGGACCACTCAGCCCTTAGAGTATATTTCCGC 2400
Db 427 AGGTCCACCTTCCAGCGCGGAGCCTAGGACCACTCAGCCCTTAGAGTATATTTCCGC 369
Qy 2401 ACTTCAGAAATTCATATCTTGGCAATCCAAAGTCCCTGCCCGCCCAATTAAGTCTCTGC 2460
Db 368 ACTTCAGAAATTCATATCTTGGCAATCCAAAGTCCCTGCCCGCCCAATTAAGTCTCTGC 309
Qy 2461 TTCCAGAAATTTGGAATCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
Db 308 -TCCAGAAATTTGGAATCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 250
Qy 2521 TCCGCGCGCGAGCTATGAGCATCTGAGCGCGCGCTTCTCTCAGCAAACTGGCGCCCGG 2580
Db 249 TCCGCGCGCGAGCTATGAGCATCTGAGCGCGCGCTTCTCTCAGCAAACTGGCGCCCGG 190
Qy 2581 ATCAGACGAGACCTCCCTTCCGACCTCTGGGAAACCTTCCAGAGGTCCAGCCCATCTCG 2640
Db 189 ATCAGACGAGACCTCCCTTCCGACCTCTGGGAAACCTTCCAGAGGTCCAGCCCATCTCG 130
Qy 2641 GAGCATCCCGAGGAAATCTCAGAGGGTGTAGGAGTGGGTGACAAGAGCTGATCTCTT 2700
Db 129 GAGCATCCCGAGGAAATCTCAGAGGGTGTAGGAGTGGGTGACAAGAGCTGATCTCTT 71
Qy 2701 CCTGTTTTGTACATAGATTTATTTTTCAGTTCCCAAGAAAGATGAATATTTTGTAAAA 2760
Db 70 CCTGTTTTGTACATAGATTTATTTTTCAGTTCCCAAGAAAGATGAATATTTTGTAAAA 11
Qy 2761 AAAAAAATAA 2770
Db 10 AAAAAAATAA 1
```

```
RESULT 9
US-09-948-941-529
; Sequence 529, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 529
; LENGTH: 11190
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-529

Query Match 45.2%; Score 1253.2; DB 5; Length 11190;
Best Local Similarity 96.2%; Pred. No. 4.3e-234;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;

Qy 1382 GGAGGGCAGTCCACACGCTGTGTGCACGTGGGTGGGCGCCACCTTTCTTGTGCAGTCCCT 1441
Db 7819 ggaatggcagcagaggtcctgtgagttccagccactggccggaaccttctctcaactt 7878
Qy 1442 AGGCGCTTTTGGGACAGGTGGCGCAGAAAGGTGCGCATAGTAGCCGAGAGGCGCCAGGC 1501
Db 7879 gctctcccactggcccaggtgcygagcaagggtgcygcatagtgcagcagagggcggagggc 7938
Qy 1502 GCGCGAGCGCCAGAGAGCCGTGGGCGAGAGAGCCCGGAAAGCGCGGCGGCGGCGCCACG 1561
Db 7939 ggcgagggccgagagcgctg9ggcgagaggaagcccg9gagggcg9g9ggggccacg 7998
Qy 1562 GCGGAGTCCAAAGCCAGAGAGGCGCGCGCCGCGCCCAAGAACGACACTGGACAAAGGCGCT 1621
Db 7999 gcggaggtccaaagcagagggcggcgccgcccgaagagccagcactggacaagggcct 8058
Qy 1622 GGGCAGCGGCGAGGTGTCAGTGTGAGGACCCCGCGGAAGCCTCTCTGGGACTGTGCTGG 1681
Db 8059 gggcaccggcgaagggtgagtgctgagggcccccccggaagcctctgggactgtcgtg 8118
Qy 1682 CACAGCGCGAGGCGCTGAAGGTGGCAGCAGCGGTGCGGAGTGGCGAGCCCGCAGCATCACC 1741
Db 8119 cacagcccgagggcctgaaggtggcagcagcagctcagtgccagcagcccgagcagcacc 8178
Qy 1742 ACGCCCGAGGGTCCAGTGTCTACTTTCCAGAGTGTGAGAGATGAAGGGCATGAAGAGAGCT 1801
Db 8179 acg9cg9aggggtccagtgctcactctccagagtgcgaagatgagggcagtgaaaggagct 8238
Qy 1802 GCTGTGTGCGCACCAAGATCAACTGAGCGGCATCAAGCTGCAAGCTGCAAGTGCACAGTCCGA 1861
Db 8239 gc9gg9tggccaccagatcaactcgagcgcatacaagctgcaactcagcgcaagtcgca 8298
Qy 1862 AGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACTACTCTGTCTTTCCTCAA 1921
Db 8299 agtgagatgaagagcagaaagtgtccacccctagtgactacacactctgtcttttcccaa 8358
Qy 1922 GCGCAGCGCAAGGCGCTCTCAACTACTGGGAGTTCGGACCGCTTGTGGGACCCAGGC 1981
Db 8359 gc9ggcagcgcaaggcctctgaactactggggactctggagcgtctgtg9ggg9cccg9gc 8418
Qy 1982 TCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCGCCCAAGGGACAGGC 2040
Db 8419 tcg9ccctagtcccccacactctgagcccatgttctgccccagcccaagggagcagggc 8478
Qy 2041 CTCACCTCTACCCAAACCCCTAGGTTCCGGTCCGGGTACAGTCTGTATCAAAACCCACGA 2100
Db 8479 ctcaactctacccaaacctcaggttgcg9gtcccg9gtcccg9gtacagctgtatcaaacccaca 8538
Qy 2101 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCGAGTCTGTAGATATAGTCTCTCTCC 2160
Db 8539 ttttctccagctcagaaacccagggcctgtgccccagtgctgagagataag9tctctctcc 8598
Qy 2161 CAGAAATCCAGCGCGGCAATGGAACCTCAGCTGGGTCTCTTAATTACAGCTTTTAAAGG 2220
Db 8599 cagaatcccgagcg9ccaaatggaacactcagctgggtcctctaataccagctctttaaagg 8658
Qy 2221 CCCAGCGCTTAGAAACCCAGCTCCTCTCGGAAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 8659 cccagcctctagaaaccccaagctcctctctcggaaccgctcactcactagagccagcagcact 8718
Qy 2281 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGGTTTCAACCTTAACCCAAAGGAGCAC 2340
```

[illegible]

QY	2101	TTTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTCGTTAGAAATATAGGTCTCTTCTCC	2160
Db	8539	ttttctccagctcagaaccacgggctctgcccacgctcgtagaataaggctctctctctcc	8598
QY	2161	CAGAATCCCAGCCGGCCCAATGAAACCTCAGCTCGGTGCTTAATTACCAGTCTTTAAAGG	2220
Db	8599	cagaa tccacagccgccaatggaaacctcaegctgggtctctaaataccagctctctaaagg	8658
QY	2221	CCCAGGCCCTAGAAACCCAAAGCTCTCTTCGGAACCGCTCACTAGAGCCAGACCAACGT	2280
Db	8659	ccagagccctagaaaccccaagctctctctcggaaacgctccacctagagccagacacacgt	8718
QY	2281	TACTCAGGGCTCTCTCCACGCTTGTAGGAGCTGAGGTTTCACGCTTAACCCAGGAGCAC	2340
Db	8719	tactcagggctcctccacgcttgtaggagctgaggtttcaacccttaacccaa -ggagcac	8777
QY	2341	AGTCCCACTCCAGCCCGGGAGCCTAGGACCCTCAGCCCTTAGGAGTATATTTCOCG	2400
Db	8778	aggtcccaactccagccc -gggagcctaggaccactcagcccttaggagtagtatctcgc	8836
QY	2401	ACTTCAGAATTCCATATCTTGGCAATCCAGCTCCCTGCCCCCAATACTTTCAGTCTCTGC	2460
Db	8837	acttcagaattccaatatcttgcgaatcccaagctcctgcccacaataaacttcagtcctgc	8896
QY	2461	TTCCAGAAATTTGGAAATCTAGTCTTCCCTCTCCTTCGTATCCGAGTCGTGGCACAACAAC	2520
Db	8897	-tccagaa tttggaaatcctagtttccctctcctcgatcccgagctcgggacacaaac	8955
QY	2521	TCCGCCCCACAGCCTATGACATCTCAGCCCGCCCTCTTCTCAGCAAACTGCGCCCGG	2580
Db	8956	tccgccccagaccta tgaagca tcc tgaagccccccctctctctgaagaaactggcccccg	9015
QY	2581	ATCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCG	2640
Db	9016	atcagagaggacctcccttcgaacctctggaaacctcccagaggctccagcccatctcg	9075
QY	2641	GAGCATCCCGAGGAAATTCGACAGGGGTATGAGAGTGGGTGACAAGAGCTGATCTCTT	2700

Query Match 17.1%; Score 473.4; DB 6; Length 519;
Best Local Similarity 99.0%; Pred. No. 9e-83;

```

? FILE REFERENCE: Aecomica-X-1
? CURRENT APPLICATION NUMBER: US/09/864,761
? CURRENT FILING DATE: 2001-05-23
? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: GB 24263.6
? PRIOR FILING DATE: 2000-10-04
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668

```

;
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5337
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000134.14
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; US-09-864-761-5337

Query Match 16.3%; Score 452; DB 7; Length 464;
Best Local Similarity 99.8%; Pred. No. 1.3e-77;
Matches 463; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1597 AGAAGCAGCACTGGACAAAGGGCTGGGACCGCCGAGGCTGAGGTGTCAGGACCCCGCC 1656
DB 464 AGAAGCAGCACTGGACAAAGGGCTGGGACCGCCGAGGCTGAGGTGTCAGGACCCCGCC 405
QY 1657 GGAAGCCTCTGGGACTGTCGCTGGGACACCGCCGAGGCTGAGGTGTCAGGACCGGCTC 1716
DB 404 GGAAGCCTCTGGGACTGTCGCTGGGACACCGCCGAGGCTGAGGTGTCAGGACCGGCTC 345
QY 1717 AGTGCCAGCACCGCAGCATCACACCGCGGAGGCTCCAGTGCTCACTTTCCAGAGTG 1776
DB 344 AGTGCCAGCACCGCAGCATCACACCGCGGAGGCTCCAGTGCTCACTTTCCAGAGTG 285
QY 1777 AGAAGATGAAGGGCATGAAGGAGTGTGTGGCCACCAAGATCAACTCGAGGCGCATCA 1836
DB 284 AGAAGATGAAGGGCATGAAGGAGTGTGTGGCCACCAAGATCAACTCGAGGCGCATCA 225
QY 1837 AGCTGCAACTCAGGCGACATGCGCAAGTGCAGATGAAGACGAAAGTGTCCACCCCTA 1896
DB 224 AGCTGCAACTCAGGCGACATGCGCAAGTGCAGATGAAGACGAAAGTGTCCACCCCTA 165
QY 1897 GTGACTACACTCTCTCTTCTCAAGCGGACGCGCAAGGCTCTGAAGTCTGAGGACT 1956
DB 164 GTGACTACACTCTCTCTTCTCAAGCGGACGCGCAAGGCTCTGAAGTCTGAGGACT 105
QY 1957 TCGGACCGCTTGTGGGACCGAGGCTCCG-CCCTAGTCCCGCCCACTCTGAGCCCATGTC 2015
DB 104 TCGGACCGCTTGTGGGACCGAGGCTCCGCGCTTAGTCCCGCCCACTCTGAGCCCATGTC 45
QY 2016 TCGGACCGCTTGTGGGACCGAGGCGCTCACTCTACCCCAACCC 2059
DB 44 TCGGACCGCTTGTGGGACCGAGGCGCTCACTCTACCCCAACCC 1

RESULT 15
US-60-278-561-1743/c

; Sequence 1743, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 1743
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; US-60-278-561-1743

Query Match 16.1%; Score 447; DB 8; Length 3534;
Best Local Similarity 96.8%; Pred. No. 1.7e-77;
Matches 456; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 86 GGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 145
DB 719 GGCC 660
QY 146 GCTGCGCTCCATCGACGACGTGGTGCGCCCTGTTGTGCGCGAGCTGGCGCGAGAGGAGCC 205
DB 659 GCTGCGCTCCATCGACGACGTGGTGCGCCCTGTTGTGCGCGAGCTGGCGCGAGAGGAGCC 600
QY 206 GGACCTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 265
DB 599 GGACCTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
QY 266 CGTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCCCGCCCGCCCGCCCGCC 325
DB 539 CGTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCCCGCCCGCCCGCCCGCC 480
QY 326 TGGGGGCTCACCTACTTTCCTGCGCGGACCTGTCTATCATCGCCCGCCCTCTATGCGCGG 385
DB 479 TGGGGGCTCACCTACTTTCCTGCGCGGACCTGTCTATCATCGCCCGCCCTCTATGCGCGG 420
QY 386 CTTACCGCGCCAGATCGAGGCGCGCTGACCTGTGCTCTATCTCGAAGAGGGGTGT 445
DB 419 CTTACCGCGCCAGATCGAGGCGCGCTGACCTGTGCTCTATCTCTCGAAGAGGGGTGT 360
QY 446 CTCAGCGCTGAGTGGTGAAGAGGTCTCCGATGTCTATGGAACAGCCTTCAGCCGCTC 505
DB 359 CTCAGCGCTGAGTGGTGAAGAGGTCTCCGATGTCTATGGAACAGCCTTCAGCCGCTC 300
QY 506 CTACTTCAAGGATCGGGCGCCACATCCAGTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 556
DB 299 CTACTTCAAGGATCGGGCGCCACATCCAGTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 249

Search completed: October 14, 2001, 08:17:47
Job time: 16272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 21:11:14 ; Search time 13115.4 Seconds
(without alignments)
3123.151 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTCTCGGACCGCCGGACC.....TGTTAAAAA.....2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues
Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US099_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US001_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US002_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US003_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US004_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US005_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US006_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US007_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US008_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US009_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US010_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US011_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2587.2	93.3	2840	49	US-60-172-360-22297	Sequence 22297, A
2	1865.2	67.3	2196	12	US-08-865-337-2	Sequence 2, Appli
3	863.8	31.2	986	12	US-08-865-337-4	Sequence 4, Appli
4	791.8	28.6	32768	53	US-60-213-178-196	Sequence 196, App
5	746.4	26.9	812	12	US-08-865-337-5	Sequence 5, Appli
6	522.2	18.8	603	20	US-09-539-800-12611	Sequence 12611, A
7	492.2	17.8	538	12	US-08-865-337-3	Sequence 3, Appli
8	488.4	17.6	639	1	PCT-US01-00663-18748	Sequence 18748, A
9	486.2	17.5	551	21	US-09-540-229-133603	Sequence 133603,
10	471.6	17.0	917	29	US-09-721-589-5669	Sequence 5669, Ap
11	470	17.0	867	29	US-09-721-589-4604	Sequence 4604, Ap
12	459.2	16.6	465	19	US-09-528-409-62595	Sequence 62595, A
13	452	16.3	464	1	PCT-US01-00663-5551	Sequence 5551, Ap
14	448	16.2	784	29	US-09-726-211-1570	Sequence 1570, Ap
15	446	16.1	2121	22	US-09-577-408-6141	Sequence 6141, Ap
16	446	16.1	2770	22	US-09-577-408-136	Sequence 136, App
17	440	15.9	479	16	US-09-205-070-29321	Sequence 29321, A
18	440	15.9	479	17	US-09-340-623-29321	Sequence 29321, A
19	440	15.9	479	17	US-09-353-690-11417	Sequence 11417, A
20	436.4	15.7	549	27	US-09-698-013-349	Sequence 349, App
21	425	15.3	461	27	US-09-698-010-4746	Sequence 4746, Ap
22	399	14.4	443	28	US-09-716-990-1013	Sequence 1013, Ap
23	393	14.2	405	17	US-09-332-782-18044	Sequence 18044, A
24	393	14.2	405	19	US-09-515-694-18044	Sequence 18044, A
25	378	13.6	406	19	US-09-528-409-92976	Sequence 92976, A
26	360.8	13.0	513	51	US-60-196-710-1772	Sequence 1772, A
27	356.6	12.9	381	24	US-09-637-086A-46608	Sequence 46608, A
28	356.6	12.9	381	25	US-09-654-617-229277	Sequence 229277,
29	356.6	12.9	381	27	US-09-684-016-229277	Sequence 229277,
30	323	11.7	447	17	US-09-306-350A-32588	Sequence 32588, A
31	318	11.5	391	16	US-09-248-797-47016	Sequence 47016, A
32	318	11.5	391	17	US-09-346-956-12130	Sequence 12130, A
33	310	11.2	389	16	US-09-233-076-11698	Sequence 11698, A
34	310	11.2	389	16	US-09-248-797-37577	Sequence 37577, A
35	310	11.2	389	17	US-09-332-782-11698	Sequence 11698, A
36	310	11.2	389	29	US-09-737-223-11698	Sequence 11698, A
37	308.4	11.1	2442	49	US-60-171-625-524	Sequence 524, App
38	308.4	11.1	2442	49	US-60-173-464-25060	Sequence 25060, A
39	308.4	11.1	2442	51	US-60-191-637-32160	Sequence 32160, A
40	308.4	11.1	2442	51	US-60-191-681-25485	Sequence 25485, A
41	308.4	11.1	2442	53	US-60-219-005-305	Sequence 305, App
42	295.6	10.7	413	12	US-08-803-609-4307	Sequence 4307, Ap
43	287.8	10.4	305	14	US-09-044-767-468	Sequence 468, App
44	287.8	10.4	305	21	US-09-540-764-23525	Sequence 23525, A
45	279.8	10.1	310	21	US-09-540-212A-61207	Sequence 61207, A


```

RESULT      1
US-60-172-360-22297
; Sequence 22297, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Deep, Dinsh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 22297
; LENGTH: 2840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2836
; OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-22297

```

Query Match 93.3%; Score 2587.2; DB 49; Length 2840;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2713; Conservative 0; Mismatches 44; Indels 10; Gaps 9;

Qy	2	GTGTCCGGAGCCGCGGACCTAGAGATCCAGAAAGCCACAGCAGCGGCGCCCGCCGCGCA	61
Db	79	gcgacctcagggcgagactctcatgtcccgacaggtccgcggtcgcgcgcggtgc	138
Qy	62	CTATTTCCAGGCTCTCGGGGCGAGGGCGCGCCACCGCCGCGCGCATGGGCTGAA	121
Db	139	ctagtt--gtggagtgtaaagctcggatgcgcgcgcacccgcgcgcctatgggtgaa	196
Qy	122	GGCGCCCCAGAAACGGTGTTCCCGTGGGCTCCATCGACAGCGTGGTGGCGCTGTTTC	181
Db	197	ggccgccagaaagacgtgttccgctgcgtccatcagacagctggtgcgctattgc	256
Qy	182	TGGCGAGCTGGCGCGAGAGAGCGGACCTGGTGCTCTTCTTGGTGCTGGGCTTCGT	241
Db	257	tgcgagctggcgagagagcgagacctggtgctccttctctgtgctgggttgcgt	316
Qy	242	GGACATTTTCTGGCTCTCAACCGCTCATCCCTACCAAGTTCCCGAGCTCACCTTCCA	301
Db	317	ggagcatctctggctgtcaacgcgcgtcatccctacaaagttcccgagctcaccttcca	376
Qy	302	GCCGAGCCGCGCCCGACCGCGCTGGCGGCCTCACCTACTTTCCCGTGGCGAGACTGTC	361
Db	377	gcccagccccgcgcgcgcgcgcgcgtggcgccctcaactatttcccgttggcgagctgtc	436
Qy	362	TATCATCGCCGCCTCTATGCCCGCTTCACGCCCCAGATCCGAGGCGCGCTCGAGCTGTC	421
Db	437	tatcatcgcgccctctaigcccgcttcacgcgccagatccgagggcgcgctcgacctgtc	496
Qy	422	CCTCTATCTTCGAGAGAGGGGGTGTCTCCAGCCGTGAGCTGGTGAGAAAGGTCTCGAGTG	481
Db	497	cctctaacctcgaagaagggtgtctccagcgtgagctgggtgaagaaggtctccgatgt	556
Qy	482	CATATGAACAGACTCAGCGGCTCTACTTCAAGGATCGGGCCACATCCAGTGCCTCTTT	541
Db	557	catatggaaagcgtcaagcgcgtccctacttcaaggatcgggcccatccagccctctt	616
Qy	542	CAGCTTCATCACAGGACCCAAATTGACAGACTCCCGTGTGGCCTTTGCTGTGGTGGGCG	601

Db	617	 cagcttcatcacaggcaccaaa	 ttgacagctcccggtg	 tggcccttgcgtggttggggc	 tggggc	676
Qy	602	CTGCCAGGCCCTGGGTCTCGGGATGCCACCTGCGCCCTGTCTGAGGATCATCTCCCTGGGT	 	 	 	661
Db	677	ctgcaggccctgggtctccgggatgtccacctgcgccctgtctgagatctcatgcttgggt	 	 	 	736
Qy	562	AGTCTTTGGGCCCATNTGGGAGCAGACAGCTGAGTCTACCTGGCAGCGCAAGGCCAACGA	 	 	 	721
Db	737	agtgtttggggcccaattggggagcagacagctgaggttcaactggcagcgcaagggaacga	 	 	 	796
Qy	722	GGACCGCAGGGCCAGACACTCAATGCCGCTGGCTGAGCGGAGCT-GGCTGCTACCTCA	 	 	 	780
Db	797	ggaccgcaggggccagacagtcactgccgggtg	 	 	 	856
Qy	781	AAGGATCATATCGCTGTGACCGCAAGATGAGGTGGCGTTTCATGGTGTGTGGCATCA	 	 	 	840
Db	857	aaggatcatactgcgtgtgacgcgaagatggaggttca	 	 	 	916
Qy	841	ACCTCTCATTTGACCTGCACACGACTCGCTGAGCTTCTGCAGCTTGCAGCTGCAGCAGAAGCTCC	 	 	 	900
Db	917	accttccattgacctgcacacgactcgtgtgagttctgagcttgcagctgcagcagaagctgc	 	 	 	976
Qy	901	TCCTGGCTGCTCTATGACCTGGGACATCTGGAAGTAGTACCCATGGCTCTAGGGAACCTGG	 	 	 	960
Db	977	cttgtgtctatgacctggacatctggaaggtaecccatgtgaccttggggaacctgg	 	 	 	1031
Qy	961	CAGATCTAGAGGAGCTGGAGCCACCCCTGGCGGCCAGACCCACATCAGCTCTTACCACA	 	 	 	1021
Db	1037	cagatctlagaggagctggagccacccttggccggcagaccacatcaacctctacaca	 	 	 	1091
Qy	1021	AGGCAITTTGCTCAGCCCAACACCTACTATTCCGGATGAACACATCTACCCCTACATCTAAC	 	 	 	1081
Db	1097	agggcata	 	 	 	1151
Qy	1081	TGCTTGCTACCACTGCTGCAACCGCAATGTGGGGGAAGCCCTGCAAGGCTTGGCGGACA	 	 	 	1141
Db	1157	tggcttgctaccactgtcgcaacgcgcaatgtgcgggaagccctgcaggccttggggcgaca	 	 	 	1211
Qy	1141	CGGCCACTGTCTCCAGGACTACAACCTACTCTCCGGGAAGACGAGGAGATCTACAGGAGCT	 	 	 	1201
Db	1217	cggccactgttcatccaggactcaactacttgcgggaagcaggagagatctcacaggagct	 	 	 	1271
Qy	1201	TCMTTGAAGTAGCCATGATGTATCCCCAACCTGCTGAAGGAGGAGCCAGCTTGCCTGG	 	 	 	1261
Db	1277	tccttgaagtagccaatgtatctcccaacctgtcgaaggagcgccagcttctgtctgg	 	 	 	1331
Qy	1261	AGGCGGGAGGAGCGCGCGGGGACCAAGCCA-GGGCACCCAGAGCCAGGTTCCCGCC	 	 	 	1311
Db	1337	aggcgggcaggagcgccggggagcaagcaggggcccagagcccgagccaaaggtctccgc	 	 	 	1391
Qy	1320	CTCCAGGACCTCATGTCCTTCGCCACCTGCTCGCATTTCTACACGGCATCTGCAAAATGG	 	 	 	1371
Db	1397	ctccaggaacctgaagtcttgcgccacctgtcgattctcagcagcgcatctgcgaatgg	 	 	 	1451
Qy	1380	GAGAGGGCAGTCCCAGCCTGTCTGACGTGGGCTGGGCCACCTTTCTTTGTGCAGTCC	 	 	 	1431
Db	1457	gagaggggagtcacacgctgtgtcagctgggttggggccaccttcttgtcagtcct	 	 	 	1511
Qy	1440	CTAGGCCGTTTTTAGGGACAGGTGCGGCACGAAGTGGCATAGTGAGCCGAGAGCGCCGAG	 	 	 	1491
Db	1517	ctaggccgcttcttgaggagacagggtgcggcagaagatggcgaatgagcgaagagccgag	 	 	 	1571
Qy	1500	CGCGCCGAGGCCGAGGAGCCGTGGGGCGAGGAAGCCCGGGAAGGCCGGCGGGGCCCA	 	 	 	1551
Db	1577	gcggccaggccgagggagccgtggggcgaaggagcccggaaggccggcgaggcgggcgcca	 	 	 	1631
Qy	1560	CGCGCGGAGTCCAAGCCAGAGAGCCCCCGCCGCCCAAGAAGCCACGACTTGACAAGGCG	 	 	 	1611
Db	1637	cggcgggagttccaaagccaggaggcccccccgccccaagaagccagcatctgacaaggggc	 	 	 	1691
Qy	1620	CTGGGCACCGGCCAGGCTGCAAGTGTACGA-CCCCCCCCGGAAGCCCTCTTGGGACTTCCG	 	 	 	1671

Db 1697 ctgggacccggcggtgagtgatgtagagaccccccccggaagcctctctgggactgtgc 1756
QY TGCCAGAGCCCGAGGCGCTGAAGGTGGCAGCACAGCGTCAAGTCCAGCAGCACCAGCAGCATC 1738
Db 1757 tggcagacccggcggtgagtgatgtagagaccccccccggaagcctctctgggactgtgc 1816
QY ACCACGCGCGGAGGTCCAGTCTCACTTCCAGAGTGAAGATGAAGGCGATGAAGGA 1798
Db 1817 accacccggcggtgagtgatgtagagaccccccccggaagcctctctgggactgtgc 1876
QY GCTGCTGGTGGCCACCAAGATCACTGAGCGCCATCAAGTCCAGTCCAGTCCAGTCCAGTCC 1858
Db 1877 gctgtgtggtggcagcagatcaactcagcgccatcaactcagcgccatcaactcagcgccat 1936
QY GCAAGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGTACTGACTACACTCTCTCTTCT 1918
Db 1937 gcaagtgcagatgaagagcagaaagtgtccacccctagtagctacactctctcttctct 1996
QY CAAGCGGAGCGCAAGGCTCTGAACACTACTGAGGACTTTCGAGCGCTTGTGGGACCCA 1978
Db 1997 caagcggcagcgcaagcctctgactactggtgagcctctggtgagcctctggtgagcct 2056
QY GGCTCGG-CCTTAGTCCCACTGTAGCCCATCTTCTGCCCCCAGCCCAAGGGGACA 2037
Db 2057 ggctcggcctctagtcaccccaactctgagccatgtctgtccccccagcccaagggaca 2116
QY GGCTCACCTCTACCCAAACCCCTAGTGTCCCGTCCGAGTACAGTCTGTATCAAAACCCA 2097
Db 2117 ggctcactctacccaaacctaggttcccggtcccgagtagctgtatcaaaccca 2176
QY CGATTTTCTCAGCTCAGAACCCAGGGCTCTGCCCGAGTGTAGAAATAGTCTCTTTC 2157
Db 2177 cgattttctcagctcagaaacccaggtctgtccccagtcgttagaataataggtctcttc 2236
QY TCCAGATCCAGCGGCGCATGTAAACCTCACGCTGGGTCTTAATACAGTCTTTAA 2217
Db 2237 tccagatcccgagcgccgaatggaacctcaactcagctgggtcttaattaccagtttaa 2296
QY AGGCCAGCCCTAGAACCCAGCTCTCTCCGGAACCGCTCACCTAGAGCAGACCAA 2277
Db 2297 aggccagccctagaaacccagctctctctcggaaacccgctcaccctagagccagaccaa 2356
QY CGTTACTCAGGCTCTCCAGCTGTAGAGCTGTAGGCTTCAACCTTAACCCAGGGAG 2337
Db 2357 cgttactcaggtctctccagctgttagagctgaggtttcacaccttaacccaa-ggag 2415
QY CACAGTCCCACTCCAGCGCGGAGCTAGGACCACTCAGCCCTAGAGTATATTC 2397
Db 2416 cacaggtcccaactccagccc-ggagcctaggaacctcagccccctagtagtatatttc 2474
QY CGCACTTCAGAAATTCATATCTTGGAAATCCCAAGCTCCCTGCCCAAAATTAACCTCAGTCC 2457
Db 2475 cgaacttcagaattccatattgtggaatccaaagctccctgcccccaataaattcagttcc 2534
QY TGCTTCCAGAAATTCGAATTCAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2517
Db 2535 tgc-tccagaatttggaaatctagtttctctctctctctctctctctctctctctctctct 2593
QY AACTCGGCCCCAGGCTATGAGCATCTGAGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2577
Db 2594 aactcgcggccagcctatgagcatcctgagccccgcctctctctctctctctctctctctctct 2653
QY CGGATCAGAGCAGGACCTCCCTTCCGACCTCTGGAACCTCCAGAGCTCCAGGCTCCAGCCATC 2637
Db 2654 cggatcagagcagggacctctctcagacctctctggaacctcccgaggggtccagccccca 2713
QY TCGGAGCATCCCGGAGGAATTCGAGAGGGGTAGGAGTGGGTGACAAAGCCCTGATCT 2697
Db 2714 tcggagcatcccgaggaaatctgcaga-gggttagagtggtgagacagagccctgatct 2772
QY CTTCTGTTTGTACATAGATTTATTTTTCAGTTCACAAAGATGATATACATTTTGTGA 2757
Db 2773 ctctcgtttgtatagatt 2832

QY 2758 AAAAAA 2764
Db 2833 aaanaaa 2839

RESULT 2
US-08-865-337-2
: Sequence 2, Application US/08865337
: GENERAL INFORMATION:
: APPLICANT: Au-Young, Janice
: APPLICANT: Covitz, Peter
: APPLICANT: Tang, Y. Tom
: APPLICANT: Murry, Lynn E.
: TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
: TITLE OF INVENTION: PROTEIN
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/865,337
: FILING DATE: Herewith
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0305 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2196 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: BRAINOT14
: CLONE: Consensus
US-08-865-337-2

Query Match 67.3%; Score 1865.2; DB 12; Length 2196; -
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2082; Conservative 7; Mismatches 50; Indels 140; Gaps 6;
QY 88 GCGCGCGCCACCGCCCGCGCCATGGGCTGAAGCGCCGCGCCAGAGAGCGCTGTCCCGC 147
Db 35 GCGCGCGCCACCGCCCGCGCCATGGGCTGAAGCGCCGCGCCAGAGAGCGCTGTCCCGC 94
QY 148 TCGGCTCCATGACAGCGTGGTGGCGCTGTGTGCTCCGAGCTGGGCGGAGAGAGCGCG 207
Db 95 TCGGCTCCATGACAGCGTGGTGGCGCTGTGTGCTCCGAGCTGGGCGGAGAGAGCGCG 154
QY 208 ACCTGGTGTCTCTTCTCTTGGTGTCTGGGCTTCTGGGAGCATTTTCTGGCTGTCAACGCG 267
Db 155 ACCTGGTGTCTCTTCTCTTGGTGTCTGGGCTTCTGGGAGCATTTTCTGGCTGTCAACGCG 214
QY 268 TCATCCTACCAACAGTTCCTCCGAGCTCACCTTCCAGCCAGCCCGCCCGCCGCGCTG 327

|||||
Db 215 TCATCCCTTACCAAGCTTCCCGAGCTACCTTCCAGCCAGCCCGCCCGACCCGCGCTG 274
Qy 328 GCGGCTCACCCTACTTCCCGTGGCGGACCTGTCTATCATCGCGCCCTCTATCCGCGCT 387
Db 275 GCGGCTCACCCTACTTCCCGTGGCGGACCTGTCTATCATCGCGCCCTCTATCCGCGCT 334
Qy 388 TCACCGCCAGATCCGAGGCGCCCTCGACCTGTCCCTCTATCTTCGAGAGGGGTGTCT 447
Db 335 TCACCGCCAGATCCGAGGCGCCCTCGACCTGTCCCTCTATCTTCGAGAGGGGTGTCT 394
Qy 448 CCAGCGGTGAGCTGGTGAAGAGTCTCCGATGTCTATGTAATGGAACAGCCTCAGCCGCTCT 507
Db 395 CCAGCGGTGAGCTGGTGAAGAGTCTCCGATGTCTCCCTCTATCTTCGAGAGGGGTGTCT 454
Qy 508 ACTTCAAGGATCGGCGCCACATTCAGTCCCTCTTCAGCTTCATCAGGCGACCAAAATGG 567
Db 455 ACTTCAAGGATYGGGCGCCACATTCAGTCCCTCTTCAGCTTCATCAGGCGACCAAAATGG 514
Qy 568 ACAGCTCCGGTGTGGCCTTTTCTGTGTGGGCGCTGCCAGGCCCTGGGTCTCCGGGATG 627
Db 515 A---CTCCGGNGTGCCTTTGCT----- 534
Qy 628 TCCACCTCGCCCTCTCTGAGGATCATGCTGGGTAGTGTGTTGGCCCAATGGGGAGCAGA 687
Db 535 -----GTGNNTGGGCCCAATGGGGAGCAGA 559
Qy 688 CAGCTGAGGTCACTTGCAGCGCAAGGCGACAGGAGCCGACGAGGCGCCACAGCTCAATG 747
Db 560 CAGCTGAGGTCACTTGCAGCGCAAGGCGACAGGAGCCGACGAGGCGCCACAGCTCAATG 619
Qy 748 CCGGTGTGGCTGAGCGAGCTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA 807
Db 620 CCGGTGTGGCTGAGCGAGCTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA 679
Qy 808 AGATGAGGTGGCTTTCATGCTGTGTCATCAACCTTCCATTTGACCTGACACCGACT 867
Db 680 AGATGAGGTGGCTTTCATGCTGTGTCATCAACCTTCCATTTGACCTGACACCGACT 739
Qy 868 CGCTGGAGCTTCTCAGCTGCAGCAGAGAGCTGCTGTGGCTCTATGACCTGGGACATC 927
Db 740 CGCTGGAGCTTCTCAGCTGCAGCAGAGAGCTGCTGTGGCTCTATGACCTGGGACATC 799
Qy 928 TGGAAAGTACCCCATGGCTTADGGAACTTGGCAGATCTAGAGGAGCTGGAGCCCAACC 987
Db 800 TGGAAAGTACCCCATGGCTTADGGAACTTGGCAGATCTAGAGGAGCTGGAGCCCAACC 859
Qy 988 CTGGCGGCGAGCCCACTACCCCTTACACAGGCGATTCCTCAGCCAGACCTACT 1047
Db 860 CTGGCGGCGAGCCCACTACCCCTTACACAGGCGATTCCTCAGCCAGACCTACT 919
Qy 1048 ATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGGCTACCACTGTGCAACCGCA 1107
Db 920 ATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGGCTACCACTGTGCAACCGCA 979
Qy 1108 ATGTGCGGGAAGCCCTCGAGGCTGGG---CGGACAGCGGCGACTGTGATCCAGGACTACA 1164
Db 980 ATGTGCGGGAAGCCCTCGAGGCTGGG---CGGACAGCGGCGACTGTGATCCAGGACTACA 1039
Qy 1165 ACTACTGCGGGAAGCAGGAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATGCA 1224
Db 1040 ACTACTGCGGGAAGCAGGAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATGCA 1099
Qy 1225 TCCCAACCTCTCTGAAGGAGCAGCCAGCTTGTGAGGCGGGGAGGAGCGCGCGGGG 1284
Db 1100 TCCCAACCTCTCTGAAGGAGCAGCCAGCTTGTGAGGCGGGGAGGAGCGCGCGGGG 1159
Qy 1285 AGCAAGCCAGGCGACCCAGAGCAAGGTTCCCGCCCTCCAGGACCCCTGAGTGTTCGCC 1344
Db 1160 AGCAAGCCAGGCGACCCAGAGCAAGGTTCCCGCCCTCCAGGACCCCTGAGTGTTCGCC 1219
Qy 1345 ACCTGCTGCGATCTACGAGCGCATCTGCAAAATGGGAGGCGGAGTCCCAAGCTGTGC 1404
|||||

Db 1220 ACCTGCTGCGATTTACGACGGCATCTCRAATGGGAGGAGGCGAGTCCCAAGCTGTGC 1279
Qy 1405 TGCAGTGGGCTGGGCGACCTTTCTTGTGCACTCCCTAGGCGGTTTGAAGGACAGTGC 1464
Db 1280 TGCAGTGGGCTGGGCGACCTTTCTTGTGCACTCCCTAGGCGGTTTGAAGGACAGTGC 1339
Qy 1465 GGCAGAAAGTGGCGATAGTGAAGCGGAGGCGGAGGCGGAGGCGGAGGAGCGTGGG 1524
Db 1340 NGCAGAAAGTGGCGATAGTGANCCGAGATGCCGAGGCTGCC----- 1380
Qy 1525 GCGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCCCAACGCGGGGAGTCCCAAGCCAGGAGC 1584
Db 1381 -----GAGNCCAAAGCCAGAGGAGC 1399
Qy 1585 CCGCGCGCCCAAGAAGCCAGCACTGGACAGGCGCTGGGCGACCGCGCAGGTCAGTGT 1644
Db 1400 CCGCGCGCCCAAGAAGCCAGCACTGGACAGGCGCTGGGCGACCGCGCAGGCGCAGTGT 1459
Qy 1645 CAGGACCCCGCCGGAAGCCCTCTCTGGGACTGTCTGGTGGCACAGCGCCGAGGCGCTGAAGTG 1704
Db 1460 CAGGACCCCGCCGGAAGCCCTCTCTGGGACTGTCTGGTGGCACAGCGCCGAGGCGCTGAAGTG 1519
Qy 1705 GCAGCAGGCTCAGGTGCCAGCACCCGCGAGCATCACCACCGCGGAGGTCAGTGTCTCA 1764
Db 1520 GCAGCAGGCTCAGGTGCCAGCACCCGCGAGCATCACCACCGCGGAGGTCAGTGTCTCA 1579
Qy 1765 CTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGGTGGCGCACCAAGATCAACT 1824
Db 1580 CTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGGTGGCGCACCAAGATCAACT 1639
Qy 1825 CGAGCGCCATCAAGCTCAACTACGCGCACAGTCCGCAAGTGCAGATGAAGAGCAGAAAG 1884
Db 1640 CGAGCGCCATCAAGCTCAACTACGCGCACAGTCCGCAAGTGCAGATGAAGAGCAGAAAG 1699
Qy 1885 TGTCCACCCCTAGTGTACTACTCTGTCTTCTCAAGCGGCGCGCAAGGCTCTGAA 1944
Db 1700 TGTCCACCCCTAGTGTACTACTCTGTCTTCTCAAGCGGCGCGCAAGGCTCTGAA 1759
Qy 1945 CTACTGGGAGCTTCCGAGCCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCCCAACTCT 2003
Db 1760 CTACTGGGAGCTTCCGAGCCGCTTGTGGGACCCAGGCTCCGCGCTTAGTCCCCCAACTCT 1819
Qy 2004 GAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGCTCAGCTCTACCCAAACCCCTAGG, 2063
Db 1820 GAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGCTCAGCTCTACCCAAACCCCTAGG 1879
Qy 2064 TTCCCGGTCCCGAGTACAGTCTGATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 2123
Db 1880 TTCCCGGTCCCGAGTACAGTCTGATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 1939
Qy 2124 GCTGTGCCCCAGTGGTTAGATATAGTCTCTTCTCCAGAAATCCAGCGGCGCAATGGA 2183
Db 1940 GCTGTGCCCCAGTGGTTAGATATAGTCTCTTCTCCAGAAATCCAGCGGCGCAATGGA 1999
Qy 2184 AACCTCAGCTGGGTCTCTAATTACCAGTCTTTAAAGGCGCAGCCCTAGAAACCCAGCT 2243
Db 2000 AACCTCAGCTGGGTCTCTAATTACCAGTCTTTAAAGGCGCAGCCCTAGAAACCCAGCT 2059
Qy 2244 CTTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAGGCTCTCTCCAGCTT 2302
Db 2060 CNTCTCGGNAACGCTACCTAGAGCCAGCAACGTTATCAGGGGTCTCTCCAGCTT 2119
Qy 2303 GTAGGAGCTGAGGTTTCCACCTTAACCCAGGAGGAGCAGGTCACGCTCCAGCCCGG 2361
Db 2120 GTAGGAGCTGAGGTTTCCACCTTAACCCAGGAGGAGGAGGTCCTCCAGCTT 2178
|||||

RESULT 3
US-08-865-337-4
; Sequence 4, Application us/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter

APPLICANT: Tang, Y. Tom
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 986 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT08
CLONE: 2809030
US-08-865-337-4

Query Match 31.2%; Score 863.8; DB 12; Length 986;
Best Local Similarity 92.9%; Pred. No. 2.4e-167;
Matches 907; Conservative 2; Mismatches 63; Indels 4; Gaps 2;
QY 668 TGGGCCCAATGGGGAGCAGACAGCTGAGGTACCTGGCAGCGCAAGGGCAAGGACCG 727
DB 1 TGGGCCCAATGGGGAGCAGACAGCTGAGGTACCTGGCAGCGCAAGGGCAAGGACCG 60
QY 728 CAGGGCCAGACAGTCAATCGCGGTGCTGCTGAGCGGAGCTGGCTACCTAAAGGATC 787
DB 61 CAGGGCCAGACAGTCAATCGCGGTGCTGCTGAGCGGAGCTGGCTACCTAAAGGATC 120
QY 788 ATACATGGCTGTGACCGCAAGATGAGGTGCGTTCATGTTGCTGTCATCAACCCCTTC 847
DB 121 ATACATGGCTGTGACCGCAAGATGAGGTGCGTTCATGTTGCTGTCATCAACCCCTTC 180
QY 848 CATTGACCTGCACACCGACTCGCTGAGGTTCTGAGCTGTCAGTGCAGAGAGCTCTGGCT 907
DB 181 CATTGACCTGCACACCGACTCGCTGAGGTTCTGAGCTGTCAGTGCAGAGAGCTCTGGCT 240
QY 908 GCTCTATGACCTGGGACATCTGGAAGGTACCCCTTGGCTTGGGACCTGGCAGATCT 967
DB 241 GCTCTATGACCTGGGACATCTGGAAGGTACCCCTTGGCTTGGGACCTGGCAGATCT 300
QY 968 AGAGGAGCTGGAGCCACCCCTGGCGCGGCGCAGACCCACTCACCTCTACCAAGGATC 1027
DB 301 AGAGGAGCTGGAGCCACCCCTGGCGCGGCGCAGACCCCTGGCTCTCAAGAGGATC 360
QY 1028 TGCCTCAGCCAAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGG 1087

DB 361 TGCCTCAGCCAAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTGCTGGCTGG 420
QY 1088 CTACCACCTGTGCGCAACCGCAATGTGCGGAAAGCCCTGCAAGGCTGG---CGGACAGCGC 1144
DB 421 CTACCACCTGTGCGCAACCGCAATGTGCGGAAAGCCCTGCAAGGCTGGGAGAGACAGCGC 480
QY 1145 CACTGTCTATCCAGGACTACAACCTACTGCGGGAAGCAGGAGAGATCTACAAGGATTTCTT 1204
DB 481 CACTGTCTATCCAGGACTACAACCTACTGCGGGAAGCAGGAGAGATCTACAAGGATTTCTT 540
QY 1205 TCAAGTAGCCCAATGATGTCATCCCAACCTGCTGAAGAGGAGCAGCAGCTTCTGGAGGC 1264
DB 541 TGAAGTAGCCCAATGATGTCATCCCAACCTGCTGAAGAGGAGCAGCAGCTTCTGGAGGC 600
QY 1265 GGGCGAGGAGCGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCCAAGGTTCCGCCCTCCA 1324
DB 601 GGGCGAGGAGCGCGCGGGGAGCAAGCCAGGCGACCCAGAGGTTCCGCCCTCCA 660
QY 1325 GGACCTGTAGTCTGCTGCGCCACCTGCTGCGATTCTACGACGCATCTGCAAAATGGAGGA 1384
DB 661 GGACCTGTAGTCTGCTGCGCCACCTGCTGCGATTCTACGACGCATCTGCAAAATGGAGGA 720
QY 1385 GGGCAGTCCACCGCTGCTGCTGCGAGTGGGCTGGGCGACCTTTCTTGTGTCAGTCCCTAGG 1444
DB 721 GGGCAGTCCACCGCTGCTGCTGCGAGTGGGCTGGGCGACCTTTCTTGTGTCAGTCCCTAGG 780
QY 1445 CGGTTTGGAGGACAGGTGCGGCGAGAGGTGCGCATAGTAGCCGAGAGCGGCGGCGGCG 1504
DB 781 CGGTTTGGAGGACAGGTGCGGCGAGAGGTGCGCATAGTAGCCGAGATGCGGAGGCTCGC 840
QY 1505 CGAGCGGAGGAGCGGTGGGCGAGAGCGCGGGAAGCGCGGCGGCGGCGGCGGCGGCGG 1564
DB 841 CGAGCGGCGGAGCGGTGGGCGAGAGCGCGGGAAGCGCGGCGGCGGCGGCGGCGGCGGCG 899
QY 1565 GGAGTCCAAAGCCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1624
DB 900 GGAGTCCAAAGCCANATGAGCNCCCGCGCANNAGCNANNAGCNANNAGCNANNAGCNCC 959
QY 1625 CACCGCGCGGAGGCTGCA 1640
DB 960 CACCGCGCGGAGGCTGCA 975

RESULT 4
US-60-213-178-196/c
; Sequence 196, Application US/60213178
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000689
; CURRENT APPLICATION NUMBER: US/60/213,178
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 1425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(32768)
; OTHER INFORMATION: n - A,T,C or G
US-60-213-178-196

Query Match 28.6%; Score 791.8; DB 53; Length 32768;
Best Local Similarity 99.2%; Pred. No. 4.2e-152;
Matches 848; Conservative 0; Mismatches 2; Indels 5; Gaps 5;
QY 1917 CTCAAGCGCGCAGCGCAAGGCGCTCTGAACTACTTGGGGAGCTTCGACCGCTTGTGGGAGCC 1976

```
Db 32768 CTAAGCGGAGGCGCAAGGCGCTCTGAACCTACTGGGAGCTTCGGACCGCTTGTTGGGAGCC 32709
QY 1977 CAGGCTCCG-CCTTAGTCCCGCCCAACTCTGAGCCCATGTTCTGCGCCCGCCAGCCCAAGGGGA 2035
Db 32708 CAGGCTCCGCGCCCTTAGTCCCGCCCAACTCTGAGCCCATGTTCTGCGCCCGCCAGCCCAAGGGGA 32649
QY 2036 CAGGCGCTCACTCTACCCAAACCCCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACC 2095
Db 32648 CAGGCGCTCACTCTACCCAAACCCCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACC 32589
QY 2096 CAGGATTTTCTCAGCTCAGAACCCAGGGGCTGTGCGCCAGTCTGTAGATATAGTCTCT 2155
Db 32588 CAGGATTTTCTCAGCTCAGAACCCAGGGGCTGTGCGCCAGTCTGTAGATATAGTCTCT 32529
QY 2156 TCTCCAGAAATCCAGCGCGGCCCAATGAAACCTCAGCTGGGTCTTAATTTACAGTCTTT 2215
Db 32528 TCTCCAGAAATCCAGCGCGGCCCAATGAAACCTCAGCTGGGTCTTAATTTACAGTCTTT 32469
QY 2216 AAAGGCCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACC 2275
Db 32468 AAAGGCCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACC 32409
QY 2276 AAGTTTACTCAGGGCTCTCTCCAGCTGTAGGAGCTGAGGTTTCACGCTTAACCCAAAGG 2335
Db 32408 AAGCTTACTCAGGGCTCTCTCCAGCTGTAGGAGCTGAGGTTTCACGCTTAACCCAA-GG 32350
QY 2336 AGCACAGGTCACCACTCCAGCGCGGAGGCTTAGGACCACTCAGCCCTCAGGAGTATATT 2395
Db 32349 AGCACAGGTCACCACTCCAGCGCGGAGGCTTAGGACCACTCAGCCCTCAGGAGTATATT 32291
QY 2396 TCCGCACTTCAGAAATCCATATCTTGGAAATCCAAAGCTCCCTGCCCAATAACTTCAGT 2455
Db 32290 TCCGCACTTCAGAAATCCATATCTTGGAAATCCAAAGCTCCCTGCCCAATAACTTCAGT 32231
QY 2456 CTTGCTTCCAGAAATTTGGAATCCTAGTTTCTCTCTCGTATCCGGAGTCTGGGACAC 2515
Db 32230 CTTGCTTCCAGAAATTTGGAATCCTAGTTTCTCTCTCGTATCCGGAGTCTGGGACAC 32172
QY 2516 AAAAATCCGCGCCCGCCAGCTATGAGCATCTGAGCGCGCGCCCTCTCTGACGAAATCTGGC 2575
Db 32171 AAAAATCCGCGCCCGCCAGCTATGAGCATCTGAGCGCGCGCCCTCTCTGACGAAATCTGGC 32112
QY 2576 CCGGATCAGAGGAGGACCTCCCTTCGAGACCTCTGCGAACTCTCCAGAGGTCAGGCCCA 2635
Db 32111 CCGGATCAGAGGAGGACCTCCCTTCGAGACCTCTGCGAACTCTCCAGAGGTCAGGCCCA 32052
QY 2636 TCTCGGAGCATCCGAGGAAATCTCAGAGGGGTAGGAGTGGGTGACAGAGCCTGAT 2695
Db 32051 TCTCGGAGCATCCGAGGAAATCTCAGAGGGGTAGGAGTGGGTGACAGAGCCTGAT 31993
QY 2696 CTCTTCTCTGTTTGTACATAGATTATTTTTCAGTTCCAGAAAGATGAATACATTTGT 2755
Db 31992 CTCTTCTCTGTTTGTACATAGATTATTTTTCAGTTCCAGAAAGATGAATACATTTGT 31933
QY 2756 TAAAAAATAATAAA 2770
Db 31932 TAAAAAATAATAAA 31918
```

```
RESULT 5
US-08-865-337-5
; Sequence 5, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT02
CLONE: J313372
US-08-865-337-5
```

```
Query Match 26.9%; Score 746.4; DB 12; Length 812;
Best Local Similarity 97.2%; Pred. No. 3.2e-143;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;

QY 1570 CCAAGCCAGGAGAGCCCCCGCCCAAGAGCCAGCAGCTGGACAGGGCTTGGGACCG 1629
Db 1 CCAAGCCAGGAGAGCCCCCGCCCAAGAGCCAGCAGCTGGACAGGGCTTGGGACCG 60
QY 1630 GCCAGGCTCAGTGTCTGAGGACCCCGGAAAGCTCTCTGGGAGTGTCTGCGTGGCAGACGCC 1689
Db 61 GCCAGGCTCAGTGTCTGAGGACCCCGGAAAGCTCTCTGGGAGTGTCTGCGTGGCAGACGCC 120
QY 1690 GAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCACCGCCG 1749
Db 121 GAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCACCGCCG 180
QY 1750 AGGCTCCAGTCTCACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGGTGG 1809
Db 181 AGGCTCCAGTCTCACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGGTGG 240
QY 1810 CCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCAGCAGTCCCAAGTGCAGA 1869
Db 241 CCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCAGCAGTCCCAAGTGCAGA 300
QY 1870 TGAAGAAGCAGAAAGTGTCCACCCCTAGTAGTACTACTCTGTCTTTTCTTCAAGCGCAGC 1929
Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTAGTACTACTCTGTCTTTTCTTCAAGCGCAGC 360
QY 1930 GCAAGGCGCTCTCAACTACTGGGAGCTTGGGAGCCGCTTGTGGGAGCCAGGCTCCG-CCT 1988
Db 361 CAAAGGCGCTCTCAACTACTGGGAGCTTGGGAGCCGCTTGTGGGAGCCAGGCTCCGCGCT 420
QY 1989 TAGTCCCGCCCAACTCTGAGCCCATGTTCTGCGCCCGCAGCCCAAGGGGACAGGCGCTCACCTC 2048
Db 421 TAGTCCCGCCCAACTCTGAGCCCATGTTCTGCGCCCGCAGCCCAAGGGGACAGGCGCTCACCTC 480
```

```
QY 2049 TACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 2108
Db 481 TACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 540
QY 2109 AGCTCAGAACCCAGGCTCTGCCCGAGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 2168
Db 541 AGCTCAGAACCCAGGCTCTGCCCGAGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 600
QY 2169 CAGCCGGCCCAATGGAAACCTCACGCTGGGTCTCTTAATTTACAGTCTTTAAAGGCCAGGCC 2228
Db 601 CAGCCGGCCCAATGGAAACCTCACGCTGGGTCTCTTAATTTACAGTCTTTAAAGGCCAGGCC 660
QY 2229 CTAGAAACCCAGCTCTCTCTCGG-AACCGCTCACCTAGAGCCAGACCAACGTTACTCAG 2287
Db 661 CTAGAAACCCAGCTCTCTCTCGGAAACCGTCACTAGAGCCAGACCAACGTTACTCAG 720
QY 2288 GGCTCTCCCGAGTCTAGAGCTAGGTTTACCCCTTAACCCAGGGAGCAGAGTCC 2347
Db 721 GGCTCTCCCGAGTCTAGAGCTAGGTTTACCCCTTAACCCAGGGAGGAAANGTCC 780
QY 2348 ACTCCAGCCCGGG 2361
Db 781 ANCTTCAGCCCGGG 794
```

RESULT 6

```
US-09-539-800-12611
; Sequence 12611, Application US/09539800
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09539,800
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: August 16, 1995
; PRIOR FILING DATE: June 27, 1994
; PRIOR FILING DATE: November 4, 1994
; PRIOR FILING DATE: October 3, 1997
; PRIOR FILING DATE: October 4, 1996
; PRIOR FILING DATE: October 4, 1997
; PRIOR FILING DATE: October 4, 1996
; PRIOR FILING DATE: October 4, 1997
; PRIOR FILING DATE: December 18, 1997
; PRIOR FILING DATE: December 20, 1996
; PRIOR FILING DATE: February 10, 1999
; PRIOR FILING DATE: February 12, 1998
; PRIOR FILING DATE: December 1, 1999
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 12611
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; OTHER INFORMATION: Incyte ID No: hu01347785
US-09-539-800-12611
Query Match 18.8%; Score 522.2; DB 20; Length 603;
Best Local Similarity 98.1%; Pred. No. 3.9e-97;
Matches 560; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 2038 GCCTCCTACCTTACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCCA 2097
Db 1 ggcctcacctctaccacaaacccctaggtctcccggtcccgagtagcagtcgtatataagtccttc 60
QY 2098 CGATTTCCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCCGTTAGATAATATAGTCTCTTC 2157
Db 61 cgatttctccagctcagaaccacaggtctgccccagtcgttagaataatagtccttc 120
QY 2158 TCCCAGAAATCCAGCCGGCCCAATGGAAACCTCACGCTGGGTCTCTTAATTTACCAAGTCTTTAA 2217
Db 121 tccagaatcccgagcgcccaatggaaacctcacgctgggtctctaattaccagtccttaa 180
QY 2218 AGCCCCAGCCCTTAGAAACCCAGGCTCTCTCGGAAACCGCTCACCTAGAGCCAGACCA 2277
Db 181 aggccagccccctagaaaccccaagctctctctcggaacccgctcacctagagccagaccaa 240
QY 2278 CGTTACTCAGGGCTCTCCAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCCAAGGGAG 2337
Db 241 cgttactcagggctctccagcttgtagagctgaggtttcaccttaacccaa-ggag 299
QY 2338 CACAGTCCCACTCCAGCCGGGGAGCCTTAGGACCACTCAGCCCTAGGAGTATATTTTC 2397
Db 300 cacaggtcccaactccagccc-gggagcctaggaaccactcagccctagagtagtatattc 358
QY 2398 CGCACTTCAGAAATTCATATCTTTCGAAATCCAAAGCTCCCTGCCCCCAAAATTCAGTCC 2457
Db 359 cgcacttcagaaattccatatcttcgaaatcccaagctccctcgccccaaataactcagtc 418
QY 2458 TGCTTCCAGAAATTTGAAATTCCTAGTCTCTCTTCCTTCCTATCCCGAGTCTGGGACACAA 2517
Db 419 tgc-tccagaatttgaaatccctagttctctctctctatcccgagctcggagacaa 477
QY 2518 AACTCGCCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCGACGAAACTGGCCC 2577
Db 478 aactcgcgcccgagcctatgagcatcctgagcccgccctctctctcctcgaacaaactggccc 537
QY 2578 CGGATCAGAGCAGGAGCTCCCTTCCTCGACCCCT 2608
Db 538 cggatcagagcaggaactcccttcggacctct 568
```

RESULT 7

```
US-08-865-337-3
; Sequence 3, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```



```

Qy 1862 AGTCAGATGAAGAGACAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCTCAA 1921
Db 82 AGTCAGATGAAGAGACAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCTCAA 23
Qy 1922 GCGGCAGCGCAAAAGGCCTCTGA 1943
Db 22 GCGGCAGCGCAAAAGGCCTCTGA 1
RESULT 9
US-09-540-229-133603
; Sequence 133603, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 133603
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01071364
; NAME/KEY: unsure
; LOCATION: 389, 423, 466, 484
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-133603

```

Query Match	17.5%	Score 486.2	DB 21	Length 551
Best Local Similarity	97.8%	Pred. No. 9	8e-90	
Matches 543	Conservative 0	Mismatches 7	Indels 5	Gaps 5
QY 2107	CCAGCTCAGAACCCAGCGGCTCTGCCCCAGTCGTTAGAAATATAGGTCTCTCTCCAGAAAT	2166		
DB 1	ccagctcagaaacca-ggctctgccccagtcgttagaata tagtctcttccccagaat	59		
QY 2167	CCCAGCCGGCCAAATGGAACCTCACGCTGGGTCCCTAATTACAGTCTTTAAAGGCCCCAGC	2226		
DB 60	cccagccggccaatggaaacctcacgctgggtcctaattaccagtcctttaaagcccacg	119		
QY 2227	CCCTAGAAACCCAAAGCTCTTCCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCA	2286		
DB 120	ccctagaaacccaagctcctctcctggaaacgctcacctagagcagagacaaacgttactca	179		
QY 2287	GGGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCTTAAACCAAGGGACACAGGTCC	2346		
DB 180	gggtctctccagcttgtagagagctgagggttccaccttaaccacaa-ggagacagggccc	238		
QY 2347	CACCTCCAGCCCGGGGAGCCTTAGGACCACCTCAGGCCCTTAGGAGTATATTTCCGCACTTCA	2406		
DB 239	caectccagccc-gggagcctaggaccactcagccccctaggatatttcgcgaactcca	297		
QY 2407	GAATTCATATCTTGCAGATCCCAAGTCCCTGCCCCCAATAAATCTCAGTCTCTGTTCCAG	2466		
DB 298	gaattccatatcttgcgaatccaagctcccttgcgcccaataaacttcagtcctgc-tccag	356		
QY 2467	AATTTGGAATPCCTAGTTTCTCTCTTCGTTATCCCGAGTCTGGGACAC-AAAACCTCGGC	2525		
DB 357	aacttggaaatccctagtttccctctccttcgtantcccgagtcctgggacacaaaaactccgcg	416		
QY 2526	CCCAGGCTATGAGCATCTCTGAGCCCGCCCTCTTCTCTGACGAACCTGAGCCCGCCGATCAG	2585		

```

Db      417  cccagncatgagatcctctgagcccccgcctcttctctgacgaaactggncccgatcag 476
QY      2586  AGCAGGACCTCCCTTCCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGGAGCA 2645
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      477  agcaggancctctccgacctctgggaacctccagagglccagcccatctcgagca 536
QY      2646  TCCCGGAGGAATCT 2660
         | ||||| |||||
Db      537  ttccggaggaaatct 551

RESULT 10
US-09-721-589-5669
; Sequence 5669, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villevall, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721,589
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167,380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5669
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-589-5669

Query Match          17.0%; Score 471.6; DB 29; Length 917;
Best Local Similarity 98.1%; Pred. No. 1.1e-86;
Matches 477; Conservative 0; Mismatches 9; Indels 0; Gaps

```

	Query Match	17.0%;	Score 471.6;	DB 29;	Length 917;
	Best Local Similarity	98.1%;	Pred. No. 1.le-86;		
	Matches 477;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps
QY	431	TCGACAGAGGGGTCTCTCAGCGCTGAGCTGTGAAGAAGGTCCTCCGATGTCATATGAA	490		
Db	3	tcgaccacgcgtccgctccagcgcgtgagctggtagaaggctcccgatgcatatgaa	62		
QY	491	CAGCCTCAGCGCTCTACTTCAAGGATCGGCCACCATCCAGTCCCTCTTCAGATTTCAT	550		
Db	63	cagcctcagcgcctcacttcaaggatcgggccacatccagtcaccttcagcttcat	122		
QY	551	CACAGGCACCAATTTGGACAGCTCCGGTGTCGCTTGTGTGGGGCTGCCAGGC	610		
Db	123	cacaggcaccaattggacagctccggctggtgcccccttgctggtgtgggctgcgaagc	182		
QY	611	CTTGGGTCTCCGGGATGTCCACCTCGCCCTGCTGAGGATCATGCTGGGTACTGTTTGG	670		
Db	183	cctgggtctccggagtgtccacctcgcccgtgtctgaggatcatgacctgggttaagtgttgg	242		
QY	671	GCCCAATGGGGAGCACAGCTGAGTCACTTGGCACGGCAAGGGCAACGAGGACCGCAG	730		
Db	243	gcccaattggggagcagacagctgaggtcaccttggcaaggcgaaggagccgcag	302		
QY	731	GGSCCAGACAGTCAATTCGGGTGTGGCTGAGCGGAGCTGGCTGTACCCTCAAAGGATCAT	790		
Db	303	gggccagacagtcaattgcgcgtgtgctgagcggagctggctgtacctgaaaggatcata	362		
QY	791	CATGCGCTGTACCGCAAGATGGAGTGGGTTCATGGTGTGTGCATCAACCCCTTCAT	850		
Db	363	catgpcgtgtgaccgcaagaagtggagggtggcgtttcatgggtgtgccaataaaccttccat	422		
QY	851	TGACCTGCACACCGACTTCGCTGGAGCTTCTGCAGCTGCACGAAGCTGCTCTGGCTGCT	910		
Db	423	tgacctgcacaccgactcgtggagcttctgcagctgcagcagaagctgctctggctgct	482		
QY	911	CTATGA	916		
Db	483	ctagga	488		

RESULT 11

US-09-721-589-4604
; Sequence 4604, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villevail, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721.589
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167.380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4604
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; FEATURE:
; LOCATION: (1)...(867)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-589-4604

Query Match 17.0%; Score 470; DB 29; Length 867;

Best Local Similarity 97.9%; Pred. No. 2.4e-86;
Matches 476; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	431	TCGAGAAGGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAGCTCTCCGATGTCATATGAA	490
Db	12	tcgcccacgcgtccgcctccacgcgtgagctggtgaagaagctccgagtcatatgaa	71
Qy	491	CAGCCTCAGCCGCTCCTACTTCAGGATCGGGCCACATCCAGTCCCTCTTCAGCTTCAT	550
Db	72	cagctcagcgcctcctacttcaaggatcgggccacatccagctccctcttcagcttcat	131
Qy	551	CACAGGACCAAAATGGACAGCTCCGGTGTGGCCTTTGCTGTGGGGCTGCCAGGC	610
Db	132	caaggcaacaaattggaagctccggtggtggtcccttctgctggtggtggtccagc	191
Qy	611	CCTGGGTCTCCGGATGCCACCTCGCCCTCTCTGAGGATCATGCCGTGGGTAGTGTGG	670
Db	192	cctgggtctccggatgtccacctcgccctgctcagagatcatgctggtggtgttgg	251
Qy	671	GCCCAATGGGAGCAGACAGCTGAGTCACTGCGCAGGCAAGGCAACGAGACCCGAG	730
Db	252	gcccaatgggagcagacagctgaggtcaacctggcagcgcaaggcaacggaccgag	311
Qy	731	GGGCAGACAGTCAATGCGGTGTGGGTGAGCGGAGCTGGCTGTACCTGAAAGGATCATA	790
Db	312	ggccagacagtcgaatccggtgtggtcagcagcgtgctgtacctgaaagatcata	371
Qy	791	CATGGCTGTGACCGCAAGATGAGGTGGCGCTTCATGTTGTGTCATCAACCCCTTCAT	850
Db	372	catgctgtgacccgaagatggaggtggcgttcatggtgtgtgccatcaaccccttcat	431
Qy	851	TGACTGCACACCCGACTCGCTGGAGCTTCTGCAGCTGCAGCAGAGCTGCTCGCTGCT	910
Db	432	tgactgcacaccgactcgtggaggttcttcagctgcagcagcagcgtcgtcctggtcgt	491
Qy	911	CTATCA 916	
Db	492	ctagga 497	

RESULT 12

US-09-528-409-62595
; Sequence 62595, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.

; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528.409
; CURRENT FILING DATE: 2000-03-17
; PRIOR FILING DATE: 60/125.453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 62595
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-62595

Query Match 16.6%; Score 459.2; DB 19; Length 465;
Best Local Similarity 99.4%; Pred. No. 3.3e-84;
Matches 461; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	638	CCTCTCTGAGGATCATGCCCTGGGTAGTGTGTTGGGCCCAATGGGAGCAGACAGCTGAGGT	697
Db	1	cctctctgagatcatcctcctgggtagttgttggcccaatgggagcagacagctgaggt	60
Qy	698	CACCTGGCAGCGCAAGGCAACGAGGACCGCAGGCGCAGACAGTCAATGCCGTGTGGC	757
Db	61	cacctggcagcgcaaggcgagggcgagggcgagggcgagggcgagggcgagggcgaggg	120
Qy	758	TGACGCGAGCTGGCTGTACCTGAAAGGATCATATGCGCTGTGACCGCAAGATGGAGGT	817
Db	121	tgagcgagctggctgtacctgaaagatcatacatgcgtgtgacggcaagatggaggt	180
Qy	818	GGCCTTACCTGTGTGCCATCAACCTTCCATTTGACCTGCACACCGACTGCTGGAGCT	877
Db	181	ggccttacctggtgtgtgccatcaaccttccattgacctgcacaccgactgcgtggagct	240
Qy	878	TCTCAGCTGCAGCAGAGCTGCTCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTA	937
Db	241	tctcagctgcagcagagctgctgctgctgctgctgctgctgctgctgctgctgctgctg	300
Qy	938	CCCCATGGCCTTAGGAACTGGCAGATCTAGAGAGCTGAGAGCCACCCCTGGCCGCC	997
Db	301	ccccatggccttaggaaacctggcagatctagaggagctggagagccacctggccgccc	360
Qy	998	AGACCCACTCACTTACCAAGGCAAGGCTGCTGACCCAAAGACCTACTATCGGATGA	1057
Db	361	agacccactcaaccttaccacaaggcattgcttcggccagagccacctactatcctgga	420
Qy	1058	ACACATACCCCTACATGTACCTGGCTGGCTACCACTGTGCGCA 1101	
Db	421	acacatcacccctacatgtacctggctggctgacctgctgacctgctgca 464	

RESULT 13

PCT-US01-00663-5551/c
; Sequence 5551, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7

FILE REFERENCE: 1600.2031-001
CURRENT APPLICATION NUMBER: US/09/736,211
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/167,843
PRIOR FILING DATE: 1999-11-29
NUMBER OF SEQ ID NOS: 1949
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1570
LENGTH: 784
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(784)
OTHER INFORMATION: n = A,T,C or G
US-09-726-211-1570

Query Match 16.3%; Score 452; DB 1; Length 464;
Best Local Similarity 99.8%; Pred. No. 1e-82;
Matches 463; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1597 AGAAGCCAGCTGGACAAAGGCGCTGGGACCGGCGGAGGTGAGTGGACACCGCCCG 1656
DB 464 AGAAGCCAGCTGGACAAAGGCGCTGGGACCGGCGGAGGTGAGTGGACACCGCCCG 405
QY 1657 GGAAGCCTCTGGAGTGTGCTGGACAGCGGCGGAGGCGCTGAGGTGCGACGAGGTC 1716
DB 404 GGAAGCCTCTGGAGTGTGCTGGACAGCGGCGGAGGCGCTGAGGTGCGACGAGGTC 345
QY 1717 AGTGCCAGCAGCCAGCATCAGCAGCGGCGGAGGCGGAGGTCAGTGTCTACTTCCAGAGTG 1776
DB 344 AGTGCCAGCAGCCAGCATCAGCAGCGGCGGAGGCGGAGGTCAGTGTCTACTTCCAGAGTG 285
QY 1777 AGAAGATGAAGGCGATGAAGGAGTGTGCTGGCCACCAAGATCAACTCGAGCGCCATCA 1836
DB 284 AGAAGATGAAGGCGATGAAGGAGTGTGCTGGCCACCAAGATCAACTCGAGCGCCATCA 225
QY 1837 AGCTGAACCTACGGCACAGTGCAGAGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTA 1896
DB 224 AGCTGAACCTACGGCACAGTGCAGAGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTA 165
QY 1897 GTGACTACACTGTCTTCTTCTCAAGCGGCGGAGGCGGAGGCTCTGAACTACTGGGACT 1956
DB 164 GTGACTACACTGTCTTCTTCTTCTCAAGCGGCGGAGGCGGAGGCTCTGAACTACTGGGACT 105
QY 1957 TCGGACCGCTGTGGGAGCCAGGCTCCG-CCTTAGTCCCGCAACTCTCAGGCCCATGTTC 2015
DB 104 TCGGACCGCTGTGGGAGCCAGGCTCCG-CCTTAGTCCCGCAACTCTCAGGCCCATGTTC 45
QY 2016 TCGCCCCAGCCCAAGGGGACAGGCGCTACCTCTACCCAAACCC 2059
DB 44 TCGCCCCAGCCCAAGGGGACAGGCGCTACCTCTACCCAAACCC 1

RESULT 14
US-09-726-211-1570
Sequence 1570, Application US/09726211
GENERAL INFORMATION:
APPLICANT: Hunter, John J.
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
THEREFOR

FILE REFERENCE: 1600.2031-001
CURRENT APPLICATION NUMBER: US/09/736,211
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 60/167,843
PRIOR FILING DATE: 1999-11-29
NUMBER OF SEQ ID NOS: 1949
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1570
LENGTH: 784
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(784)
OTHER INFORMATION: n = A,T,C or G
US-09-726-211-1570

Query Match 16.2%; Score 448; DB 29; Length 784;
Best Local Similarity 97.3%; Pred. No. 7.8e-82;
Matches 467; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 80 GGGCAGGGCGCGCGCCACCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 139
DB 43 GCGCGAGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 102
QY 140 GTTCCCGCTGCGCTCCATCGACGACGTGCTGCGCCCTGTTTGTGCGAGCTGGCGCGAGA 199
DB 103 GTTCCCGCTGCGCTCCATCGACGACGTGCTGCGCCCTGTTTGTGCGAGCTGGCGCGAGA 162
QY 200 GGAGCGCGACCTGGTGTCTCTTTCCTTGTGCTGGGCTTCTGTGGAGCATTTTGTGCTGT 259
DB 163 GGAGCGCGACCTGGTGTCTCTTTCCTTGTGCTGGGCTTCTGTGGAGCATTTTGTGCTGT 222
QY 260 GAAACGCGTCACTCCCTACCAAGCTTCCCGAGCTCACCTTCCAGCC---CAGCCCCCGCCC 316
DB 223 CAACCGCTCACTCCCTACCAAGCTTCCCGAGCTCACCTTCCAGCC---CAGCCCCCGCCC 282
QY 317 CGACCGCGCTGGCGGCGCTCACCTACTTTCCGTCGCGGCGACCTGTCTATCATCGCGCCCT 376
DB 283 CGACCGCGCTGGCGGCGCTCACCTACTTTCCGTCGCGGCGACCTGTCTATCATCGCGCCCT 342
QY 377 CTAATGCGCGCTTCAACCGCGCGAGATCCGAGCGCGCGCTGCGACCTGTCTATCATCGCGCG 436
DB 343 CTAATGCGCGCTTCAACCGCGCGAGATCCGAGCGCGCGCTGCGACCTGTCTATCATCGCGCG 402
QY 437 AGGGGTGTCTCCAGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB 403 AGGGGTGTCTCCAGCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 462
QY 497 CAGCGCGCTCTACTTCAAGGATCGGGCGCCACATCCAGTCCCTTTCAGCTTCAATCAGG 556
DB 463 CAGCGCGCTCTACTTCAAGGATCGGGCGCCACATCCAGTCCCTTTCAGCTTCAATCAGG 522

RESULT 15
US-09-577-408-6141/c
Sequence 6141, Application US/09577408
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Tillinghast, John
APPLICANT: Sinku, Ankura
APPLICANT: Liu, Chenghua
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
Polypeptides
FILE REFERENCE: 792
CURRENT APPLICATION NUMBER: US/09/577,408
CURRENT FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 8502
SOFTWARE: pc_genes Version 1.0
SEQ ID NO 6141
LENGTH: 2121

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1003)...(1386)
; OTHER INFORMATION: similar to g13002527 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-577-408-6141

Query Match 16.18; Score 446; DB 22; Length 2121;
Best Local Similarity 100.0%; Pred. No. 2.7e-81;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 ATGGGGCTGAAGCGCGCCAGAACAGCTGTTCCTCGCTCCATCGACGACGTGGTG 170
Db 2121 ATGGGGCTGAAGCGCGCCAGAACAGCTGTTCCTCGCTCCATCGACGACGTGGTG 2062

Qy 171 CGCTGTGTTGCTGCCGAGCTGGCGCGAGAGCGGACCTGGTGCTCCTTCCCTGGTG 230
Db 2061 CGCTGTGTTGCTGCCGAGCTGGCGCGAGAGCGGACCTGGTGCTCCTTCCCTGGTG 2002

Qy 231 CTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGTTCCCGAG 290
Db 2001 CTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGTTCCCGAG 1942

Qy 291 CTCACCTTCCAGCCAGCCCGCGCCGAGCCCGCTGGCGGCTCAGCTACTTTCCCGTG 350
Db 1941 CTCACCTTCCAGCCAGCCCGCGCCGAGCCCGCTGGCGGCTCAGCTACTTTCCCGTG 1882

Qy 351 GCCGACCTGTCTATCATCGCGCGCCCTCTATGCCCGCTTCAACCGCCAGATCCGAGCGCC 410
Db 1881 GCCGACCTGTCTATCATCGCGCGCCCTCTATGCCCGCTTCAACCGCCAGATCCGAGCGCC 1822

Qy 411 GTCGACCTGTCCCTTATCTCGAGAGGGGGTGTCTCCAGCCGTGAGTGGTGAAGAAG 470
Db 1821 GTCGACCTGTCCCTTATCTCGAGAGGGGGTGTCTCCAGCCGTGAGTGGTGAAGAAG 1762

Qy 471 GTCCTCCGATGTATGGAACAGCGCTCAGCGCTTCTACTTCAAGGATCGGGCCACATC 530
Db 1761 GTCCTCCGATGTATGGAACAGCGCTCAGCGCTTCTACTTCAAGGATCGGGCCACATC 1702

Qy 531 CAGTCCCTTCTCAGCTTCATCACAGG 556
Db 1701 CAGTCCCTTCTCAGCTTCATCACAGG 1676

Search completed: October 14, 2001, 07:39:50
Job time: 37716 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 20:00:54 ; Search time 218.79 Seconds
(without alignments)
2398.512 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTGTCGGAGCCGCGACC.....TGTTAAAAA.....AAAAA 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/lna/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1865.2	67.3	2196	2	US-08-865-337A-2
2	863.8	31.2	986	2	US-08-865-337A-4
3	746.4	26.9	812	2	US-08-865-337A-5
4	492.2	17.8	538	2	US-08-865-337A-3
5	62.4	2.3	7218	1	US-08-232-463-14
6	57.2	2.1	1931	2	US-09-130-114-2
7	55.4	2.0	7218	1	US-08-232-463-14
8	52.6	1.9	38506	3	US-09-320-878-19
9	47.6	1.7	30001	1	US-08-125-468-1
10	47.6	1.7	30001	2	US-08-474-933-1
11	45	1.6	2370	1	US-08-104-072B-7
12	45	1.6	2370	1	US-08-351-413-8
13	45	1.6	2370	2	US-09-025-583-8
14	44.8	1.6	2219	3	US-08-510-646B-17
15	44.2	1.6	1146	1	US-08-482-385A-1
16	44.2	1.6	1780	2	US-08-933-821-5
17	44.2	1.6	1780	3	US-08-960-507-5
18	44.2	1.6	2728	1	US-08-482-385A-5
19	44.2	1.6	4524	2	US-08-845-998-7
20	44.2	1.6	4524	3	US-09-206-537-7
21	43.6	1.6	2580	3	US-09-050-863-2
22	43.6	1.6	2582	2	US-08-480-994-6
23	43.6	1.6	2582	2	US-08-616-844-6
24	43.6	1.6	2582	2	US-08-599-654-6
25	43.6	1.6	2582	2	US-08-485-573-6
26	43.6	1.6	2582	3	US-08-944-868A-6
27	43.6	1.6	2582	3	US-08-944-423A-6

28	43.6	1.6	2582	3	US-08-925-743-6	Sequence 6, Appl
29	43.6	1.6	2582	3	US-08-944-496-6	Sequence 6, Appl
30	43.6	1.6	2582	4	US-08-925-767-6	Sequence 6, Appl
31	43.6	1.6	5452	2	US-09-130-114-1	Sequence 1, Appl
32	43.6	1.6	9600	4	US-08-910-647-1	Sequence 1, Appl
33	43.6	1.6	10596	1	US-07-884-811-15	Sequence 15, Appl
34	43.6	1.6	10596	1	US-07-885-971-15	Sequence 15, Appl
35	43.6	1.6	10596	1	US-08-087-783A-15	Sequence 15, Appl
36	43.6	1.6	10596	1	US-08-194-088B-15	Sequence 15, Appl
37	43.6	1.6	10596	2	US-08-194-087-15	Sequence 15, Appl
38	43.6	1.6	10596	5	PT-US933-04648-15	Sequence 15, Appl
39	43.2	1.6	1722	4	US-09-385-028-15	Sequence 15, Appl
40	43.2	1.6	11604	4	US-09-385-028-13	Sequence 13, Appl
41	43.2	1.6	15079	4	US-09-385-028-1	Sequence 1, Appl
42	43	1.6	2712	3	US-09-025-691-4	Sequence 4, Appl
43	43	1.6	12001	1	US-08-458-568A-11	Sequence 11, Appl
44	42	1.5	409	1	US-07-626-618A-8	Sequence 8, Appl
45	42	1.5	409	1	US-08-333-977-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-865-337A-2
; Sequence 2, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Kovitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
; US-08-865-337A-2

Query Match 67.3%; Score 1865.2; DB 2; Length 2196;									
Best Local Similarity 91.4%; Pred. No. 0;									
Matches 2082; Conservative 7; Mismatches 50; Indels 140; Gaps 6;									
QY	88	GC	GC	GC	GC	GC	GC	GC	GC
DB	35	GC	GC	GC	GC	GC	GC	GC	GC
QY	148	TG	GC	GC	GC	GC	GC	GC	GC
DB	95	TG	GC	GC	GC	GC	GC	GC	GC
QY	208	AC	TC	GC	GC	GC	GC	GC	GC
DB	155	AC	TC	GC	GC	GC	GC	GC	GC
QY	268	TC	AT	CC	TC	AC	CA	AG	TC
DB	215	TC	AT	CC	TC	AC	CA	AG	TC
QY	328	GC	GC	GC	GC	GC	GC	GC	GC
DB	275	GC	GC	GC	GC	GC	GC	GC	GC
QY	388	TC	AC	GC	GC	GC	GC	GC	GC
DB	335	TC	AC	GC	GC	GC	GC	GC	GC
QY	448	CC	AG	CC	TC	AG	TC	GC	AT
DB	395	CC	AG	CC	TC	AG	TC	GC	AT
QY	508	AC	TC	CA	AG	TC	GC	AT	TC
DB	455	AC	TC	CA	AG	TC	GC	AT	TC
QY	568	AC	AG	TC	GC	AT	TC	GC	AT
DB	515	A	---	CT	CG	NG	TG	GC	TT
QY	628	TC	CA	CT	GC	GC	CT	GC	TC
DB	535	---	---	---	---	---	---	---	---
QY	688	CA	GC	TC	AG	TC	GC	AT	TC
DB	560	CA	GC	TC	AG	TC	GC	AT	TC
QY	748	CC	GC	TC	AG	TC	GC	AT	TC
DB	620	CC	GC	TC	AG	TC	GC	AT	TC
QY	808	AG	TC	GC	AT	TC	GC	AT	TC
DB	680	AG	TC	GC	AT	TC	GC	AT	TC
QY	868	CG	TC	GC	AT	TC	GC	AT	TC
DB	740	CG	TC	GC	AT	TC	GC	AT	TC
QY	928	TG	GC	AT	TC	GC	AT	TC	GC
DB	800	TG	GC	AT	TC	GC	AT	TC	GC
QY	988	CT	GC	AT	TC	GC	AT	TC	GC
DB	860	CT	GC	AT	TC	GC	AT	TC	GC
QY	1048	AT	CG	GC	AT	TC	GC	AT	TC
DB	920	AT	CG	GC	AT	TC	GC	AT	TC

QY	1108	AT	TC	GC	GC	GC	GC	GC	GC
DB	980	AT	TC	GC	GC	GC	GC	GC	GC
QY	1165	AC	TC	GC	GC	GC	GC	GC	GC
DB	1040	AC	TC	GC	GC	GC	GC	GC	GC
QY	1225	TC	CC	CA	AC	TC	GC	GC	GC
DB	1100	TC	CC	CA	AC	TC	GC	GC	GC
QY	1285	AG	CA	AA	GC	GC	GC	GC	GC
DB	1160	AG	CA	AA	GC	GC	GC	GC	GC
QY	1345	AC	TC	GC	GC	GC	GC	GC	GC
DB	1220	AC	TC	GC	GC	GC	GC	GC	GC
QY	1405	TG	CA	GC	TC	GC	GC	GC	GC
DB	1280	TG	CA	GC	TC	GC	GC	GC	GC
QY	1465	GC	CA	AA	GC	GC	GC	GC	GC
DB	1340	GC	CA	AA	GC	GC	GC	GC	GC
QY	1525	GC	CA	AA	GC	GC	GC	GC	GC
DB	1381	---	---	---	---	---	---	---	---
QY	1585	CC	GC	GC	GC	GC	GC	GC	GC
DB	1400	CC	GC	GC	GC	GC	GC	GC	GC
QY	1645	CAG	AC	CC	CC	CC	CC	CC	CC
DB	1460	CAG	AC	CC	CC	CC	CC	CC	CC
QY	1705	GC	AG	CA	CG	TC	GC	GC	GC
DB	1520	GC	AG	CA	CG	TC	GC	GC	GC
QY	1765	CT	TC	CA	GC	TC	GC	GC	GC
DB	1580	CT	TC	CA	GC	TC	GC	GC	GC
QY	1825	CG	AG	CC	CA	TC	GC	GC	GC
DB	1640	CG	AG	CC	CA	TC	GC	GC	GC
QY	1885	TG	TC	CA	CC	TC	GC	GC	GC
DB	1700	TG	TC	CA	CC	TC	GC	GC	GC
QY	1945	CT	AC	TC	GC	GC	GC	GC	GC
DB	1760	CT	AC	TC	GC	GC	GC	GC	GC
QY	2004	GAG	CC	CA	TC	GC	GC	GC	GC
DB	1820	GAG	CC	CA	TC	GC	GC	GC	GC
QY	2064	TT	CC	GC	TC	GC	GC	GC	GC
DB	1880	TT	CC	GC	TC	GC	GC	GC	GC
QY	2124	GC	TC	GC	GC	GC	GC	GC	GC
DB	1940	GC	TC	GC	GC	GC	GC	GC	GC
QY	2184	AAC	TC	CA	GC	TC	GC	GC	GC

Db 2000 AACCTCACGCTGGTCTTAATTACCAAGTCTTTAAAGGCCAGCCCTAGAAACCAAGCT 2059
QY 2244 CTTCTCGG-AACCGCTACCTAGAGCCAGACCAAGCTTACTCAGGCTCCTCCAGCTT 2302
Db 2060 CNTCTCGGNAACCGTACCTAGAGCCAGACCAAGCTTAYTCAGGGGTCCTNCCAGCTT 2119
QY 2303 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGGACACAGGTCCTCCACCTCCAGCCCGGG 2361
Db 2120 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGGAGGAAAGGTTCCCANCTTCAGCCCGGG 2178

RESULT 2

US-08-865-337A-4
; Sequence 4, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT08
; CLONE: 2809030
US-08-865-337A-4

Query Match 31.2%; Score 863.8; DB 2; Length 986;
Best Local Similarity 92.9%; Pred. No. 1.9e-198;
Matches 907; Conservative 2; Mismatches 63; Indels 4; Gaps 2;
QY 668 TGGGCCCAATGGGGAGCAGACAGCTGAGTGCACCTGGCAGCGCAAGGCAACGAGGACCG 727
Db 1 TGGGCCCAATGGGGAGCAGACAGCTGAGTGCACCTGGCAGCGCAAGGCAACGAGGACCG 60
QY 728 CAGGGCCACAGACGTCAATCGCGGTGGCTGAGCGGAGCTGGCTACTGAAAGATC 787

Db 61 CAGGGCCACAGACAGTCAATGCCGCTGTGGCTGAGCGGAGCTGGCTGTACC:GAAGGATC 120
QY 788 ATACATGCGCTGTGACCGCAAGATGGAGTGGCTTCATCGTGTGTCCTCAACACCTTC 847
Db 121 ATACATGCGCTGTGACCGCAAGATGGAGTGGCTTCATCGTGTGTCCTCAACACCTTC 180
QY 848 CATTGACCTGTCACACCGACCTCGCTGGAGCTTCTGCAGCTGCAGCAGAAGTGTCTCTGGCT 907
Db 181 CATTGACCTGTCACACCGACCTCGCTGGAGCTTCTGCAGCTGCAGCAGAAGTGTCTCTGGCT 240
QY 908 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGGCAGATCT 967
Db 241 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGGCAGATCT 300
QY 968 AGAGGAGCTGGAGCCACCCCTGCGCGGCAGACCCACTCACCTCTACACAGGSCAT 1027
Db 301 AGAGGAGCTGGAGCCACCCCTGCGCGGCAGACCCCTGCGCGGCCTGCTRCACAGGSCAT 360
QY 1028 TGCTTCAGCCCAAGACCTTACTATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGG 1087
Db 361 TGCTTCAGCCCAAGACCTTACTATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGG 420
QY 1088 CTACCACTGTGCAACCGCAATGTGCGGAAAGCCCTGCGAGCCCTGGG---CGGACACGGC 1144
Db 421 CTACCACTGTGCAACCGCAATGTGCGGAAAGCCCTGCGAGCCCTGCGAGACAGACGGC 480
QY 1145 CACTGTCTACCCAGGACTACCACTACTGCGCGGAGAGAGGAGATCTACAGGAGTTCTT 1204
Db 481 CACTGTCTACCCAGGACTACCACTACTGCGCGGAGAGAGGAGATCTACAGGAGTTCTT 540
QY 1205 TGAAGTAGCCAATGATGTCATCCCAACCTGCTGAAGGAGCGGAGCCAGCTTGTCTGGAGGC 1284
Db 541 TGAAGTAGCCAATGATGTCATCCCAACCTGCTGAAGGAGCGGAGCCAGCTTGTCTGGAGGC 600
QY 1265 GCGGAGGAGCGCGCGGGGAGCAAGCCAGGGACCCAGAGCCCAAGGTTCCGCGCTCCA 1324
Db 601 GCGGAGGAGCGCGCGGGGAGCAAGCCAGGGACCCAGAGCCCAAGGTTCCGCGCTCCA 660
QY 1325 GGACCTGAGTGTCTGCGCCACCTGCTGCATTTCTAGCAGGATCTGCAAAATGGGAGGA 1384
Db 661 GGACCTGAGTGTCTGCGCCACCTGCTGCATTTCTAGCAGGATCTGCAAAATGGGAGGA 720
QY 1385 GGCAGTCCCACGCTGTGCTGCACGCTGGCTGGCCACCTTTCTTGTGAGTCCCTTAGG 1444
Db 721 GGCAGTCCCACGCTGTGCTGCACGCTGGCTGGCCACCTTTCTTGTGAGTCCCTTAGG 780
QY 1445 CCGTTTGGAGGACAGTGGCGGAGAGGTGCGCATAGTAGCGGAGAGCGGCGGCGG 1504
Db 781 CCGTTTGGAGGACAGTGGCGGAGAGGTGCGCATAGTAGCGGAGATGCCGAGGCTGC 840
QY 1505 CGAGGCGGAGGCGCTGGGGCGAGGAAGCCCGGAGGCGCGGCGGCGGCGGCGGCGGCGG 1564
Db 841 CGAGGCGGAGGCGCTGGGGCGAGGAAGCCCGGAGGCGCGGCGGCGGCGGCGGCGGCGG 899
QY 1565 GGAGTCCAAGCCAGAGAGCGCCCGCGCCCAAGAGAGCAGCTGGGAGAGGCGCTGGG 1624
Db 900 GGAGTCCAAGCCANATGAGCNCNCNCCGNCNANNANGACNAGGNCNCTTGGG 959
QY 1625 CACCGCGGAGGTGCA 1640
Db 960 CACCGCGGAGGTGCA 975

RESULT 3
US-08-865-337A-5
; Sequence 5, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE

; TITLE OF INVENTION: 1 PROTEIN
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/865,337A
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION NUMBER:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0305 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 812 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: BLADTUT02
 ; CLONE: 1313372
 ; US-08-865-337A-5

Query Match 26.9%; Score 746.4; DB 2; Length 812;
 Best Local Similarity 97.2%; Pred. No. 2.7e-170;
 Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;

QY	1570	CCAAGCCAGAGGAGCCCGCCGCGCCCAAGAGCCAGCAGCTGGACAGGGGCTGGGACCG	1629
DB	1	CCAAGCCAGAGGAGCCCGCCGCGCCCAAGAGCCAGCAGCTGGACAGGGGCTGGGACCG	60
QY	1630	GCCAGGGTGCGAGTGTACAGGACCCCGCGGAAAGCTCTCTGGGACTGTGCGTGGCACAGCCC	1689
DB	61	GCCAGGGCGCAGTGTACAGGACCCCGCGGAAAGCTCTCTGGGACTGTGCGTGGCACAGCCC	120
QY	1690	GAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCGCGGACATCACCCGCCGG	1749
DB	121	GAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCGCGGACATCACCCGCCGG	180
QY	1750	AGGTCCTCAGTGTCTACATTTCCAGAGTGCAGAGATGAAGGGCATGAAGGAGCTGTGTTGG	1809
DB	181	AGGTCCTCAGTGTCTACATTTCCAGAGTGCAGAGATGAAGGGCATGAAGGAGCTGTGTTGG	240
QY	1810	CCACCAAGATCAACTCGAGGCGCCATCAAGTGCACATCACGGCACATGCGAAGTGCAGA	1869
DB	241	CCACCAAGATCAACTCGAGGCGCCATCAAGTGCACATCACGGCACATGCGAAGTGCAGA	300
QY	1870	TGAAGAGCAGAAAGTGTCCACCCCTAGTACACTGTCTTTCTTCTCAAGGGGCGAG	1929
DB	301	TGAAGAGCAGAAAGTGTCCACCCCTAGTACACTGTCTTTCTTCTCAAGGGGCGAG	360
QY	1930	GCAAGGCCCTCTGAACCTACTGTGGGACTTCGGACCGCTGTGGGACCCAGCGCTCCG-CCT	1988
DB	361	CARAAGGCCCTCTGAACCTACTGTGGGACTTCGGACCGCTGTGGGACCCAGCGCTCCG-CCT	420

QY	1989	TAGTCCCCCAACTCTCAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCACCTC	2048
DB	421	TAGTCCCCCAACTCTCAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCACCTC	480
QY	2049	TACCCAAACCCCTAGTGTCCGGTCCCGGAGTACAGTCTGTATCAAAACCCAGATTTTCTCC	2108
DB	481	TACCCAAACCCCTAGTGTCCGGTCCCGGAGTACAGTCTGTATCAAAACCCAGATTTTCTCC	540
QY	2109	AGCTCAGAACCCAGGAGCTGTGCCCCAGTGGTGTAGATATAGGTCTCTTCCAGATCC	2168
DB	541	AGCTCAGAACCCAGGAGCTGTGCCCCAGTGGTGTAGATATAGGTCTCTTCCAGATCC	600
QY	2169	CAGCCGGCCCAATGGAACCTCAGCTGGTCTTAATACCACTCTTTAAAGGCCAGCCCC	2228
DB	601	CAGCCGGCCCAATGGAACCTCAGCTGGTCTTAATACCACTCTTTAAAGGCCAGCCCC	660
QY	2229	CTAGAAACCCAGCTCTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAG	2287
DB	661	CTAGAAACCCAGCTCTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAG	720
QY	2288	GGTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCAAGGGAGCAGGTTCC	2347
DB	721	GGTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCAAGGGAGGAAAGGTCC	780
QY	2348	ACCTCCAGCCCGG 2361	
DB	781	ANCTTCAGCCCGG 794	

RESULT 4
 US-08-865-337A-3
 ; Sequence 3, Application US/08865337A
 ; Patent No. 5972649
 ; GENERAL INFORMATION:
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Covitz, Peter
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Murry, Lynn E.
 ; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
 ; TITLE OF INVENTION: 1 PROTEIN
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/865,337A
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0305 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 538 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single


```
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match      2.1%; Score 57.2; DB 2; Length 1931;
Best Local Similarity 43.6%; Pred. No. 0.00014;
Matches 257; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 50 CCGGGCCCGGCACTATTTCAGGCTGTGGGGGAGGGGGCCGCCACCGCCGCGGC 109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 ccccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 110 CATGGGCTGAAGGCCCGCCAGAGAGCGCTGTCCCGCTCGGCTCCATCCAGCAGTGGT 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 cctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 GCGGCTGTTGCTGCGGAGCTGGGCGGAGAGCGGAGCTGGTGTCTCTTCTTGGT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 cgtctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 GTGGGCTTCGTGGAGCATTTTGTGGCTGTCAACCGGCTCATCCCTACCAACGTTCCGA 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 ccccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgct 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 GCTACCTTCCAGCCAGCCCGCCCGCCAGCCCGCTGGCGGCTCACTTCTCCCGT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 cctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgct 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GCGGCACTGTCTATCATCGCGGCTCTATGCGCGCTTACCGCCCGCAGATCCGAGGCGC 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 cctcgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 410 GTGCGACTGCTCTATCTCTCGAGAAGGGGTGTCTCCAGCGGTGAGTGGTGAGAA 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 ccccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 470 GGTCTCGATGTATATGAACAGCTCAGCGCTCTCTACTTCAAGGATCGGCGCCACAT 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 cctcgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 530 CCAAGTCCCTTTCAGCTATCAAGCAGCACCACAAATGAGACAGTCCGGTGTGCGCTTGC 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 ccccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 590 TGTGGTGGGCGCTCCAGGCGCTGTCTCCGGGATGTCACCTCGGCC 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 ccccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcccgctcc 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)883-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14
```

```
Query Match      2.0%; Score 55.4; DB 1; Length 7218;
Best Local Similarity 4.3%; Pred. No. 0.00065;
Matches 17; Conservative 220; Mismatches 156; Indels 0; Gaps 0;

QY 1453 AGGACAGGTGCGCAGAGGTGCGCATAGTACGAGCGAGAGCGCGGCGCGGCGCG 1512
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1513 AGGAGCGGTGGGCGGAGAGAGCGCGGAGCGCGGCGCGGCGCGGCGCGGAGTCCA 1572
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1573 AGCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1632
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1633 AGGTGCGAGTGTGAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1692
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1693 GCCTGAAGTGGCAGCAGCGCTCAGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGG 1752
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1753 GTCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGTCGCCA 1812
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1813 CCAAGATCAACTCGAGCGCGCATCAAGCTGCAAC 1845
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1074 RRRRRRRATCGCAAGCTCCTCGACCTGCGAGC 1042
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 8
US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
```


Db 1875 GCGGGGGCGCGTGGCGGGGTGTCAGTGGCGGACAGTGTGATGAGCTGATGGCGCGTGC 1934
QY 225 TTGGTCTGGGCTTCGTGGAGCATTTTCTGGGTGTCAACCGGGTGTATCCCTACCAACGTT 284
Db 1935 ATGGCTACCTCGCGGGCGCGCGGGATGACGCCCTACGGCATCTGCTGGACAGCCTC 1994
QY 285 CCGAGCTCACCTTTCAGCCAGCCCGCGCGGAGCCCGCTGGGGCCCTCACCTACTTT 344
Db 1995 GCGGTCTCAACGGATGTCGCGCGCGCGCGGACCGGTCGCGCAACTGCGTC 2054
QY 345 CCGTGGCGGACCTGTCTATCATCGCGCGCTCTATGCGCGCTTACCGCCAGATCCGA 404
Db 2055 AGGAGCGCGCGCGGCTTCCCGCGGTGACGCTTCCCGCGCTCCGCC---TCCCC 2111
QY 405 GCGCGCTGACCTGTCCCTC 425
Db 2112 GCGCGCTGCGCTCTCCATC 2132

RESULT 12
US-08-351-413-8
; Sequence 8, Application US/08351413
; Patent No. 5750867
; GENERAL INFORMATION:
; APPLICANT: Williams, Mark
; APPLICANT: Leemans, Jan
; TITLE OF INVENTION: Maintenance of male-sterile plants
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 2046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/351,413
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/899,072
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/970,849
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-102PCT
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Oryza sativa
; STRAIN: Akihikari
; FEATURE:
; NAME/KEY:
; LOCATION: 1..1808

OTHER INFORMATION: /label= PT42
OTHER INFORMATION: /note= "sequence comprising anther specific promoter PT42"
FEATURE:
NAME/KEY:
LOCATION: 1748..1755
OTHER INFORMATION: /label= TATA
OTHER INFORMATION: /note= "TATA Box"
FEATURE:
NAME/KEY:
LOCATION: 1780
OTHER INFORMATION: /note= "transcription initiation site determined by primer extension"
FEATURE:
NAME/KEY:
LOCATION: 1809
OTHER INFORMATION: /label= ATG
OTHER INFORMATION: /note= "ATG start of translation of rice '42 gene"
US-08-351-413-8

Query Match 1.6%; Score 45; DB 1; Length 2370;
Best Local Similarity 51.0%; Pred. No. 0.13; Mismatches 125; Indels 3; Gaps 1;
Matches 133; Conservative 0;

QY 165 GTGGTGGCGCTGTTCGTGGAGTGGCGGAGAGGAGCGGACCTGTGCTCTCTTTC 224
Db 1875 GCGGGGGCGCGTGGCGGGGTGTCAGTGGCGGACGATGATGATGGCGCGTGC 1934
QY 225 TTGGTCTGGGCTTCGTGGAGCATTTTCTGGGTGTCAACCGGCTTACCTACCAACGTT 284
Db 1935 ATGGCTACCTCGCGGGCGCGCGGGATGACGCCCTACGGCATCTGCTGGACAGCCTC 1994
QY 285 CCGAGCTCACCTTTCAGCCAGCCCGCGCGGAGCCCGCTGGGGCCCTCACCTACTTT 344
Db 1995 GCGGTCTCAACGGATGTCGCGCGCGCGGACCGGTCGCGCAACTGCGTC 2054
QY 345 CCGTGGCGGACCTGTCTATCATCGCGCGCTCTATGCGCGCTTACCGCCAGATCCGA 404
Db 2055 AGGAGCGCGCGCGGCTTCCCGCGGTGACGCTTCCCGCGCTCCGCC---TCCCC 2111
QY 405 GCGCGCTGACCTGTCCCTC 425
Db 2112 GCGCGCTGCGCTCTCCATC 2132

RESULT 13
US-09-025-583-8
; Sequence 8, Application US/09025583
; Patent No. 5977433
; GENERAL INFORMATION:
; APPLICANT: Williams, Mark
; APPLICANT: Leemans, Jan
; TITLE OF INVENTION: Maintenance of male-sterile plants;
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 2046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,583
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/351,413

Search completed: October 14, 2001, 03:51:57
Job time: 28263 sec

Query Match 1.68; Score 44.2; DB 1; Length 1146;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 21:01:20 ; Search time 644.85 Seconds
(without alignments)
2699.144 Million cell updates/sec

Title: US-09-380-337-1

Perfect score: 2772

Sequence: 1 GGTGTCGGAGCCGCGACC.....TGTAAAAA..... 2772

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*

```
1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2772	100.0	2772	19 AAV57415	Tumour suppressor
2	1865.2	67.3	2196	20 AAV80659	Human multiple end
3	1312.2	47.3	9180	19 AAV57416	Tumour suppressor
4	863.8	31.2	986	20 AAV80661	Human multiple end
5	746.4	26.9	812	20 AAV80662	Human multiple end
6	492.2	17.8	538	20 AAV80660	Human multiple end
7	234	8.4	253	20 AAV86611	EST clone BE3. Ho
8	66	2.4	67	16 AAT22851	Human gene signatu
9	52.6	1.9	13842	21 AAZ87297	S. venezuelae macr
10	52.6	1.9	36778	21 AAZ87318	S. venezuelae pik
11	52.6	1.9	37948	21 AAZ87285	S. venezuelae pik

12	52.6	1.9	38506	21	AAV575633	Nucleotide sequenc
13	52.6	1.9	38506	21	AAV56001	Recombinant cosmid
14	52.2	1.9	10732	21	AAA10594	Gene encoding a su
15	52.2	1.9	15872	18	AAT68715	Streptomyces venez
16	52.2	1.9	15872	21	AAZ87283	S. venezuelae vep
c 17	50.8	1.8	114955	20	AAZ53491	Human adenosine Al
18	49.6	1.8	1725	21	AAA10503	Trehalose-releasin
19	49.6	1.8	2218	21	AAA10516	Human cytoskeletal
20	48.8	1.8	7962	21	AAZ58980	Sorghum bicolor (L
21	48.2	1.7	1929	19	AAV57472	Human adenosine Al
22	48	1.7	114955	20	AAZ53491	Nucleotide sequenc
c 23	47.8	1.7	5857	21	AAA58471	Total DNA sequence
c 24	47.6	1.7	30001	18	AAT61016	S. aureofaciens DN
c 25	47.6	1.7	30001	20	AAZ05110	Gene encoding a su
c 26	47	1.7	10732	21	AAA10594	HSV-2 strain SB5 C
c 27	47	1.7	21034	19	AAV62154	HSV-2 strain SB5 C
c 28	47	1.7	117213	19	AAV62176	Nucleotide sequenc
29	45.8	1.7	18660	21	AAA58472	Human type VI aden
30	45.4	1.6	4942	20	AAZ00461	Human G protein-co
31	45.2	1.6	1128	21	AAA30619	DNA encoding human
32	45.2	1.6	1128	21	AAA30724	Human CEPR (hCEPR)
33	45.2	1.6	1648	21	AAZ11700	PEG-1 gene. Homo
34	45.2	1.6	2776	18	AAT73366	Sequence comprisin
35	45	1.6	2370	15	AAQ53880	S. clavuligerus cl
36	44.8	1.6	15079	16	AAQ91580	Streptomyces prist
37	44.2	1.6	1065	17	AAT58554	S. avermitilis BCK
38	44.2	1.6	1146	16	AAO83793	Human NEW angiopo
39	44.2	1.6	1413	22	AAF26792	Human TIE ligand N
40	44.2	1.6	1780	20	AAZ36343	Human PRO178 CDNA
41	44.2	1.6	1780	21	AAZ77516	Human anglogenesis
42	44.2	1.6	1780	22	AAC97375	S. avermitilis bkd
43	44.2	1.6	2728	16	AAO83797	Streptomyces prist
44	44.2	1.6	4496	17	AAT58553	Nucleotide sequenc
c 45	44.2	1.6	4524	20	AAV33912	

ALIGNMENTS

```
RESULT 1
AAV57415
ID AAV57415 standard; cDNA; 2772 BP.
AC
XX AAV57415;
XX
DT 11-JAN-1999 (first entry)
XX
DE Tumour suppressor gene MEN1 transcribed sequence (cDNA).
XX
KW Familial multiple endocrine neoplasia type 1; PMEN1; MEN1;
KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 111..1943
FT polyA_signal 2744..2748
FT /*tag= a
FT /*tag= b
XX
PN WO9839439-A1.
XX
PD 11-SEP-1998.
XX
PF 04-MAR-1998; 98WO-US04258.
XX
PR 05-MAR-1997; 97US-0040269.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS,
PI Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;
```


Qy	2701	CCTGTTTGTACATAGATTATTTTCCTCAGAAGAAGATGAATCATTTTTGTTAAAA	2760
Db	2701	cctgtttgtacatagatttatttttcgaagaagtgaatacatttgtaa	2760
Qy	2761	AAAAAAAAAAAAA	2772
Db	2761	aaaaaaaaaaaaa	2772
RESULT	2		
AAV80659			
ID	AAV80659	standard; cDNA; 2196 BP.	
XX			
AC	AAV80659;		
XX			
DT	15-MAR-1999	(first entry)	
DE	Human multiple endocrine receptor neoplasia type 1 protein cDNA.		
XX			
KW	Human; multiple endocrine neoplasia type 1 protein; MEND-1;		
KW	tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;		
KW	detection; diagnosis; drug screening; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	58..1758	
FT	/product=	"multiple endocrine neoplasia type 1 protein"	
FT	/transl_except=	(pos:466..468,aa:Xaa)	
FT	/transl_except=	(pos:539..540,aa:Xaa)	
FT	/transl_except=	(pos:877..879,aa:Xaa)	
FT	/transl_except=	(pos:991..993,aa:Xaa)	
FT	/transl_except=	(pos:1306..1308,aa:Xaa)	
FT	/transl_except=	(pos:1318..1320,aa:Xaa)	
FT	/transl_except=	(pos:1327..1329,aa:Xaa)	
FT	/transl_except=	(pos:1339..1341,aa:Xaa)	
FT	/transl_except=	(pos:1360..1362,aa:Xaa)	
FT	/transl_except=	(pos:1384..1386,aa:Xaa)	
FT	/note=	"Xaa is unspecified"	
XX			
PN	WO9854324-A1.		
XX			
PD	03-DEC-1998.		
XX			
PF	29-MAY-1998;	98WO-US10957.	
XX			
PR	29-MAY-1997;	97US-0865337.	
XX			
PA	(INCY-) INCYTE PHARM INC.		
XX			
PI	Au-Young J, Covitz PA, Murry LE, Tang YT;		
XX			
DR	WPI; 1999-059839/05.		
DR	P-PSDB; AAW86348.		
XX			
PT	New isolated multiple endocrine neoplasia type 1 protein - used to		
PT	develop products for the diagnosis, treatment and prevention of		
PT	cancers and multiple endocrine neoplasia.		
XX			
PS	Claim 5; Fig 1; 67pp; English.		
XX			
CC	The present sequence encodes human multiple endocrine neoplasia type 1		
CC	protein (MEND-1). MEND-1 plays a role in multiple endocrine neoplasia		
CC	when one or both normal genetic copies of MEND-1 are mutated and no		
CC	longer able to suppress tumorigenesis. MEND-1 can be used for treating		
CC	or preventing cancers and multiple endocrine neoplasia. MEND-1		
CC	polynucleotides can also be used for gene therapy. Products from the		
CC	present invention can also be used for detection, diagnosis and drug		
XX			
SQ	Sequence 2196 BP; 445 A; 712 C; 623 G; 385 T; 31 other;		

[illegible]

Qy	1108	ATGTGCGGGGAAGCCCTGCAGGCCCTGGG----	CGGACACAGCCCACTTCTATCACCAGGACTACA	1164
Db	980	: : : : : : :	atgtcggggaasccctgcaggccttgcagagacgagcactcttcacgacctaccagactaca	1039
Qy	1165	ACTACTCCCGGGGAAGACGAGGAGATCTCAAGGAGTTCTTTGAAGTCAGCCAAATGATGTC	1224	
Db	1040	: : : : : : : :	actactgcgggaagaacgagagatactaagaagtctcttgaaagtagccaattgatgca	1099
Qy	1225	TCCCCAACCTGCTTAAGAGGAGCACGAGCTTGCGAGGCGGCGGAGGAGCGCGCGGGGG	1284	
Db	1100	tccccaaaccttgtctgaaggaaggaagcagcagcttgtctggagggcgagagagcgagcgaggggg	1159	
Qy	1285	AGCAAAGCCAGGGCACCCAGAGCCAAGTTCCGCCCTCCAGGACCCTGAGTGCTTCGCGCC	1344	
Db	1160	agcaaagccaggcaccccagagccaaggttcgcgcctccaggaeccctgagtgctctcgccc	1219	
Qy	1345	ACCTGCTGCCATTCTACGACGGCATCTGCAATGGGAGGAGGGGACAGTCCCAGCGCTGTGC	1404	
Db	1220	accttgtcgatatctacgacggcatctgcaaatggagagggcgagtcgcccgccctgtgc	1279	
Qy	1405	TGCAGTGGGCTGGGCACCTTTCTTGTGCAGTGCCTAGCGCGCTTTTGAGGACACAGGTGC	1464	
Db	1280	tgcactgagggctggggcacctttctgtcnagatccccatangccgttttgangggacaggtgc	1339	
Qy	1465	GGCAGAAGTGTCCCATAGTGAGCGGAGAGCGGCGCGGCGGCGGAGCGGAGCGCGTGGG	1524	
Db	1340	ngcagaaggtgtgcatagtgancagagatgtccgaggtctgc-----	1380	
Qy	1525	GCGAAGAGCCCGGGGAAGCGCGCGGGGGGCCACGCGGGGAGTGCANATCACAGGAGC	1584	
Db	1381	-----gagncacaagccagaggagc	1399	
Qy	1585	CCCCGCGCCCAAGAGCCAGCAGCTGGAAGGGGCTGGGCACCGCCGAGGTCGCAGTGT	1644	
Db	1400	ccccgcgcgcaagaagcagcaacttggaacaaaggcctgggcaacggcgcaaggcgagtggt	1459	
Qy	1645	CAGGACCCCCCGAAGCCCTCTTGGGACATGTGCGTGGACACAGCCGCGAGCGCCCTGAAGGTG	1704	
Db	1460	caggagccccccggaagcctctgggaactgtgcgtggcaacggcgaagccctgaaggtgt	1519	
Qy	1705	GCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATATCACACC CGCGGAGGGTCCAGTGTCTN	1764	
Db	1520	gcagcagcggctcagggtgtccagcaccgcagcatccaacgcgcgaggggtccagtgctca	1579	
Qy	1765	CTTTTCCAGCTGAGAGATGAAGGGCATCAAGGAGCTGCTGTGTGGCCACCACGAATCAACT	1824	
Db	1580	ctttccagagtgaagaagaagaagggaatgaaggagctgtgtgtggccaccaagaaccaact	1639	
Qy	1825	CGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCAAGTCAGATGAAGAAGCAGAAG	1884	
Db	1640	cgagcgcca tcaagctgcaactcacgcgcagcagtcgcaagtgcagatgaagaagcagaaga	1699	
Qy	1885	TGTCACCCCTAGTGACTACACTCTCTTTTCTCAAGGGGAGCGCACGAAGGCCTCTCGAA	1944	
Db	1700	tgtccacccctagtgtactacactctgtcttctccaaagcgcgaggaagagcgctctgtga	1759	
Qy	1945	CTACTGGGGGACTTCGACCGCTTGTGGGAGCCAGGCTCCG--CCTTAGTTCGCCCAACTCT	2003	
Db	1760	ctactggggagacttcgagccgcttgtgggacccaggtccgcctcttagtcccccaactctt	1819	
Qy	2004	GAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGCGCTCACCTCTACCCAAACCTTAGG	2063	
Db	1820	gagcccaatgttctgcocccagcccaaaaggagacagggcctcaacctta cccaaaaccttagg	1879	
Qy	2064	TTCCCGGTCCCGAGTACAGTCTGTATCAACGCCACGATTTTCTCGAGCTCAGAACCCAGG	2123	
Db	1880	ttcccggttcccgagtacagtctgtata tcaaacccaagattttctecagctcagaaacccag	1939	
Qy	2124	GCTGTGCCCCAGTCTAGATAATAGTCTCTCTCCAGATCCACCGCGCCCAATCGA	2183	
Db	1940	gctctgccccagtcggttagaata tagtgtctcttccccagaatccccagccggcmcatgga	1999	
Qy	2184	AACCTCACGCTGGGTCCCTAATTAACCACTCTTTAAAGCGCCAGCGCCCTAGAAACCCCAAGCT	2243	

```

|||||
Db 2000 aaactcacgtggtcctaataaccagtccttaagagccagcccttagaaacmmaagct 2059
QY 2244 CCTCCTCGG-AACCGCTCACCTAGAGCCAGACCAACAGTTACTCAGGGCTCCTCCAGCTT 2302
Db 2060 cntctcgnaaccytcacctagagccagacacgttaytcagggtctcctncagctt 2119
QY 2303 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGCAGACAGGTCCTCCACCTCCACGCCGGG 2361
Db 2120 gtaggagctgaggtttacccttnaccctnaccacaaaggagaaangtcccancttagcccggg 2178

RESULT 3
AAV57416
ID AAV57416 standard; cDNA: 9180 BP.
XX
AC AAV57416;
XX
DT 11-JAN-1999 (first entry)
XX
DE Tumour suppressor gene MEN1.
XX
KW Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;
KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human; ds.
XX
OS Homo sapiens.
XX
FH Key
FT CDS 2288..8060 Location/Qualifiers
FT FT /*tag= a
FT FT /*note= "contains introns"
FT FT 1680..1766
FT FT /*tag= b
FT FT /*number= 1
FT FT 167..2264
FT FT /*tag= c
FT FT /*number= 1
FT FT 2265..2732
FT FT /*tag= d
FT FT /*number= 2
FT FT 2733..4296
FT FT /*tag= e
FT FT /*number= 2
FT FT 4297..4505
FT FT /*tag= f
FT FT /*number= 3
FT FT 4506..4715
FT FT /*tag= g
FT FT /*number= 3
FT FT 4716..4844
FT FT /*tag= h
FT FT /*number= 4
FT FT 4845..5176
FT FT /*tag= i
FT FT /*number= 4
FT FT 5177..5217
FT FT /*tag= j
FT FT /*number= 5
FT FT 5218..5297
FT FT /*tag= k
FT FT /*number= 5
FT FT 5298..5385
FT FT /*tag= l
FT FT /*number= 6
FT FT 5386..6024
FT FT /*tag= m
FT FT /*number= 6
FT FT 6025..6161
FT FT /*tag= n
FT FT /*number= 7
FT FT 6162..6622
FT FT /*tag= o

```

```

FT exon /number= 7
FT FT 6623..6758
FT FT /*tag= p
FT FT /*number= 8
FT FT 6759..7195
FT FT /*tag= q
FT FT /*number= 8
FT FT 7196..7360
FT FT /*tag= r
FT FT /*number= 9
FT FT 7361..7577
FT FT /*tag= s
FT FT /*number= 9
FT FT 7578..9180
FT FT /*tag= t
FT FT /*number= 10
FT FT 8861..8865
FT FT /*tag= e
FT FT
FT FT
XX WO9839439-A1.
XX
XX 11-SEP-1998.
PD
XX 04-MAR-1998; 98WO-US04258.
XX
XX 05-MAR-1997; 97US-0040269.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
XX Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;
XX Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
XX
XX WPI; 1998-506360/43.
XX P-PSDB; AAW29749.
XX
XX Protein and gene associated with multiple endocrine neoplasia type 1
XX - useful in gene therapy and to diagnose sufferers of, and those
XX susceptible to, this condition by detecting protein absence or gene
XX mutation(s)
XX
XX Claim 32; Page 57-62; 75pp; English.
XX
XX This is the nucleotide sequence of human MEN1, a novel tumour
XX suppressor gene which is associated with multiple endocrine
XX neoplasia type 1. MEN1 genomic DNA can be obtained from genomic
XX libraries using MEN1-specific probes and PCR primers. A cDNA clone
XX (see AAV57415) is also provided. MEN1 encodes a 67.5 kDa protein
XX (see AAW28749), termed menin, that exhibits no apparent similarities
XX to previously known proteins. The lack of a functional menin
XX polypeptide, either by absence of the protein, its alteration
XX and/or associated mutations in the MEN1 gene, have been identified
XX in individuals with familial multiple endocrine neoplasia type 1
XX (FMEN1) and suffering from multiple endocrine neoplasia type 1.
XX The identification of MEN1 provides a new window into the mechanism
XX of endocrine tumorigenesis, facilitates accurate early diagnosis of
XX MEN1 associated cancers, and provides preclinical identification of
XX individuals with the FMEN1 syndrome, i.e. cancer-free individuals
XX that are at high risk of acquiring MEN1 associated tumours. It
XX also provides a basis for gene therapy. The MEN1 nucleic acids may
XX also be used therapeutically to produce menin recombinantly and to
XX produce transgenic animals useful in research.
XX
XX Sequence 9180 BP; 1936 A; 2604 C; 2612 G; 2027 T; 1 other;
SQ

```

```

Query Match 47.3%; Score 1312.2; DB 19; Length 9180;
Best Local Similarity 96.5%; Pred. No. 3.2e-283;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 1382 GGAGGGCAGTCCACGCCCTGTGTGCACCTGGGGCAGCTTTCCTGTGAGTCCT 1441
Db 7499 ggatggccagagcagggtcctgagttccagcactggccggaaccttgctcacctt 7558

```

QY	1442	AGGCCGTTT	TAGGGACAGGTGCGGAGAAAGTGTCGCATAGTGAGCCGAGAGCCCGAGGC	1501
		I	I	
Db	7559	gctctcccaactg	gccagagtgcggcagaaagtgcgcatagtgagcgcagagcgccagagc	7618
QY	1502	GCCGAGGCCGAGG	AGCCGTGGGCGAGGAACCCCGGGAAGCCGCGCGCGCGGCCGCCACG	1561
Db	7619	ggcggagcgcgag	gagccgtggggcgagaaagcccggaagcgcggcgggcgggccacg	7678
QY	1562	GCGGGAGTCC	AAGCCAGAGAGGCCCGCGGCCCAAGAGCCAGCACTGAGCAAGAGGCCT	1621
		I	I	
Db	7679	gcggggagtgcc	aagccagagagagcccccgcggcccaagagccagcactgacaaagggcct	7738
QY	1622	GGCACCCGCCAG	GGTGCAGTGTCAAGACCCCCCGGAAGCCTCCTGGACATGTCGTGG	1681
Db	7739	ggcacccgcgcag	gtgcagtgccagcccccccgaagcctcctggactgtcgctgg	7798
QY	1682	CACAGCCGAGG	GGCTGAAGTGGCAGACAGCGCTCAGGTGCACAGCACCCGACGAGTCACC	1741
Db	7799	cacagcccgag	gcctggaagtggcagcacggtcagtgccagcacccgcagcatcacc	7858
QY	1742	ACCGCGGAGGT	CTCAGTCTCAGTTCACAGTGCAGAAAGATGAAGGCGATGAAGAGCT	1801
Db	7859	accgcggaggg	tccagtgctcacttccagagtgagaagatgaaggagcatgaaagagct	7918
QY	1802	GCTGTGGCCAC	CAAGATCAATCGAGGCGCCATCAAGCTGCAACTACGGCACAGTCGCA	1861
Db	7919	gctgggtggcc	accagatcaactcgagcgccatcaagctgcaactcacggcagctgcga	7978
QY	1862	AGTGCAGATGA	GAGACAGAAAGTCTCCACCCTAGTGACTACACTCTGTCTTCCTCAA	1921
Db	7979	agtgcgagtgaa	gaagcagaaagtgccaccctagtgaactacactgtctcttccctcaa	8038
QY	1922	GCGCAGCGCAA	AGGCGCTCTGAACACTGCGGAGCTTCGGACCGCTGTGGGACCCACGC	1981
Db	8039	gcggcagcgaa	agcctgaactactgggagcttcgacgcttgtgggaccacagc	8098
QY	1982	TCGCGCTTAGT	TCCCCCACTGTGAGCCCATGTTGCGCCCCCAGGCCCAAGGGGACAGGCC	2041
Db	8099	tccgctctagt	cccccaactctgagcccatgtctgcgccagcccaaggggacagcc	8158
QY	2042	TCACCTCTAC	CCAAACCCCTAGGTTCCGCTCCGAGTAGTACAGCTGTATCAACCCACGAT	2101
Db	8159	tccctctacca	aaacccctctaggttccgggtcccgagtcagctctgatacaaacccacgat	8218
QY	2102	TTTCTCCAGT	TCAGAACCCAGGGCTCTGCCAGTCGTAGAAATATAGGTCTCTTCTCCC	2161
Db	8219	ttctccagct	cagaacccagggctgtgccagctgttagaataataggctctctctccc	8278
QY	2162	AGAAATCCCA	CGGCCCAATGGAACCTCACGCTGGGTCTTAATTACAGAGTCTTTAAAGGC	2221
Db	8279	agaaatccag	cggcccaatggaaacctcacgtgggtctcaatcacagctctttaaaggc	8338
QY	2222	CCAGCCCTAGA	AAACCAAGCTCTCTCGGAACCGCTCACCTAGAGCGACAGCAACGTT	2281
Db	8339	ccagccctaga	aaacccaagctctctctctcgaaacgcgtcaactagagccagaccacagctt	8398
QY	2282	ACTCAGGGCT	CTCTCCAGCTGTAGAGCTGAGGTTTCACCTTTAACCCCAAGGGAGCACA	2341
Db	8399	actcagggtc	ctctccagctgtgtaggagctgaggttcaaccttaacccaaggagacaca	8458
QY	2342	GGTCCCACT	CCAGCCGGGAGCTAGGACACTACGCCCTAGGAGTATATTTCCGCA	2401
Db	8459	gggtcccaact	ccagccggggagcttaggaacctcagccctcagagatatttccgcga	8518
QY	2402	CTTTCAGAA	TTCATCTTTCGGAATCCAGGCTCCCTGCCCCCAATAACTTCAGTCTGCT	2461
Db	8519	ctttcagaat	tccaatactctggaatccaagctccctgcgcccaataacttcagtccttgct	8578
QY	2462	TCAGAAATTT	GGAATCTCTAGTTTCTCTCTCTGCTATCCCGAGTCTGGGACACAAAAC	2521
Db	8579	tccagaatttg	gaattccctagtttctctctctgataccccagctctgggacacaaact	8638

Qy	2532	CCGCCCCCGCCCTATGAGCATCTGTAGCCCCCGCCCTCTTCTCTGAGAACTGGCCCGCGGA	2538
Db	8639	ccgccccccagcctatgagcatctgagcccgccctcttctctgacgaaactggccccgga	8698
Qy	2582	TCAGAGCAGGACCTCCCTTCCGACCCCTCTGGGAACCTCCACAGAGTCCAGGCCCACTCGG	2641
Db	8699	tcagagcagagactccctccgacctctgggaacctcccagaggtccagctcatctcgg	8758
Qy	2642	AGCATCCCGGAGGAAATCTGCAGAGGGGTTAGGAGTGGGTGACAGAGCCCTCATCTCTTC	2701
Db	8759	agcatcccgaggaaatctgcagaggggttagagagtgggtgacaagagcctgatctcttc	8818
Qy	2702	CTGTTTGTACATAGATTATTTTCAGTTCCCAAGAAAGATGAATACATTTTGTAAAAA	2761
Db	8819	ctgtttgtacatagattatttttcagttcccaagaagatgaatcacattttgttaaaa	8878
Qy	2762	AAAAAAAAAA 2770	
Db	8879	aaatataaa 8887	
RESULT 4			
ID	AAV80661		
XX	AAV80661 standard; cDNA; 986 BP.		
AC	AAV80661;		
DT	15-MAR-1999 (first entry)		
DE	Human multiple endocrine receptor neoplasia type 1 clone 2809030.		
XX	Human; multiple endocrine neoplasia type 1 protein; MEND-1;		
KW	tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;		
KW	detection; diagnosis; drug screening; ss.		
XX	Homo sapiens.		
XX	W09854324-A1.		
PD	03-DEC-1998.		
XX	29-MAY-1998; 98WO-US10957.		
XX	29-MAY-1997; 97US-0865337.		
PA	(INCY-) INCYTE PHARM INC.		
PI	Au-Young J, Covitz PA, Murry LE, Tang YT;		
DR	WPI; 1999-059839/05.		
PT	New isolated multiple endocrine neoplasia type 1 protein - used to		
PT	develop products for the diagnosis, treatment and prevention of		
PT	cancers and multiple endocrine neoplasia.		
XX	Disclosure; Page 48; 67pp; English.		
CC	The present sequence represents Incyte clone 2809030 which is used to		
CC	make up the consensus human multiple endocrine neoplasia type 1 protein		
CC	(MEND-1) (see AAV80659). MEND-1 plays a role in multiple endocrine		
CC	neoplasia when one or both normal genetic copies of MEND-1 are mutated		
CC	and no longer able to suppress tumorigenesis. MEND-1 can be used for		
CC	treating or preventing cancers and multiple endocrine neoplasia. MEND-		
CC	polynucleotides can also be used for gene therapy. Products from the		
CC	present invention can also be used for detection, diagnosis and drug		
CC	screening.		
XX	Sequence 986 BP; 206 A; 284 C; 305 G; 154 T; 37 other;		

Query Match 31.2%; Score 863.8; DB 20; Length 986;
Best Local Similarity 92.9%; Pred. No. 1.7e-183;
Matches 907; Conservative 2; Mismatches 63; Indels 4;

```
QY 668 TGGGCCCAATGGGAGCAGACAGCTGAGGTACCTGGCAGCGCAAGGCAACGAGGACCG 727
Db 1 tggggccaatggggagacagacagctaggtaccttggcagcgcaagggcaacgagagacg 60
QY 728 CAGGGCCGACAGACATCAATCCGCGGTGGCTGAGCGGAGCTGCTACTGAAAGGATC 787
Db 61 cagggccgacagatcaatgccgggtggctgagcgagctgctactgaaagatc 120
QY 788 ATACATGCGCTGACCGCAAGATGAGGTGGGTTTCATGGTGTGNGCCATCAACCCCTTC 847
Db 121 atacaatgctgtgacccgcaagatggaggtggcgttcatgggtgtgcatcaacccttc 180
QY 848 CATTTGACCTGCACACGAGCTCGGTGAGGTCTTGTGACGTGCAGCAGAAAGCTGCTTGGGT 907
Db 181 cattgacctgcacacgactgctgagcttctgagctgcagacagagctgcttggct 240
QY 908 GCTCTATGACCTGGGACATCTGAAAGGTACCCCATGGCTTAGGGAACCTGGCAGATCT 967
Db 241 gctctatgacctgggacatctgaaagggtaccctatggccttaggaaacctggcagatct 300
QY 968 AGAGGAGCTGGAGCCACCCCTGGCGGCGCAGACCCACTCACCTTACCAAGGCGAT 1027
Db 301 aagaggctggagccacccctggcggccagaccsrgggtgctgctrcacaaggagcat 360
QY 1028 TGCTTCAGCCAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGG 1087
Db 361 tgcttcagccaagacctactatcggtatgaacacatctaccctacatgacctggtgct 420
QY 1088 CTACCACTGTGCAACCGCAATGTGCGGGAAGCCCTGACAGCGCTGGG---CGGACACGCC 1144
Db 421 ctaccactgtcgcaacgcaatgtgcggaasccttcgagggccctgfcagagacacgagc 480
QY 1145 CACTGTATCCAGGACTACACTACTGTCGGAAGAGGAGAGATCTACAAGAGTCTT 1204
Db 481 cactctatccaggactacactacttcgcggaagcagggagatctacaaggagtctt 540
QY 1205 TGAAGTAGCAATGATGTATCCCAACCTGCTGAAGAGGAGGAGGAGCTGTGTGGAGGC 1264
Db 541 tgaagtagcaaatgagtcatcccaacctgtgagaggagcagcagcttctgtaggagc 600
QY 1265 GGGCGAGGAGCGGCGGGGAGCAAGAGCGGACCCAGGACCCAGGTTCCGCCCTCCA 1324
Db 601 gggcgaggagcgggcggggagcaagcagggcagccagagccaaagggttcgcgacctca 660
QY 1325 GGACCTGTAGTGTCTGCCACCTGCTGCGATTCTACGAGCGCATCTGCAATGGGAGGA 1384
Db 661 ggacctgagtgcttcgcccacctgctgcgattctacgacgcatctgcaaatgggagga 720
QY 1385 GGGCAGTCCACACGCTGTGTGCACGTGGGCTGGGCGACCTTTTCTTGTGAGTCCCTAGG 1444
Db 721 gggcagtcacacgctgtgtgcacgtgggctgggctgggcaacctttctgtngcagtcctang 780
QY 1445 CGCTTTTGGAGGACAGGTGCGGAGAGGTGCGCATAGTGAGCCGAGAGCGGCGGCGG 1504
Db 781 cegttttgagagacaggtgngcagaaggtgcgcatagtgancggagatgcccagggctgc 840
QY 1505 CGAGGCGGAGGAGCCGTGGGGCGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCGGCGG 1564
Db 841 cgaggcc-angagccgtgngcttagaaagcccggnnangccgcnngnngggcccaangca 899
QY 1565 GGAGTCCAGGACAGGAGCGCCCGCCCAAGAGAGGACGACTGGACAAAGGCGCTGGG 1624
Db 900 ggaagtcgaagccaanagagcnnccnccgcncaannagcnannangagcnaggncccttggg 959
QY 1625 CACCGCCGAGGGTGCA 1640
Db 960 cncggccagngtga 975
RESULT 5
AAV80662
ID AAV80662 standard; cdna; 812 BP.
```

```
XX AAV80662;
XX AC
XX DT
XX DE Human multiple endocrine receptor neoplasia type 1 clone 1313372.
XX KW Human; multiple endocrine neoplasia type 1 protein; MEND-1;
XX KW tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
XX KW detection; diagnosis; drug screening; ss.
XX OS Homo sapiens.
XX PN W09854324-A1.
XX PD 03-DEC-1998.
XX PF 29-MAY-1998; 98WO-US10957.
XX PR 29-MAY-1997; 97US-0865337.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Au-Young J, Covitz PA, Murry LE, Tang YT;
XX WPI; 1999-059839/05.
XX DR
XX PT New isolated multiple endocrine neoplasia type 1 protein - used to
XX PT develop products for the diagnosis, treatment and prevention of
XX PT cancers and multiple endocrine neoplasia.
XX PS Disclosure; Page 49; 67pp; English.
XX CC The present sequence represents Incyte clone 1313372 which is used to
XX CC make up the consensus human multiple endocrine neoplasia type 1 protein
XX CC (MEND-1) (see AAV80659). MEND-1 plays a role in multiple endocrine
XX CC neoplasia when one or both normal genetic copies of MEND-1 are mutated
XX CC and no longer able to suppress tumorigenesis. MEND-1 can be used for
XX CC treating or preventing cancers and multiple endocrine neoplasia. MEND-1
XX CC polynucleotides can also be used for gene therapy. Products from the
XX CC present invention can also be used for detection, diagnosis and drug
XX CC screening.
XX SQ Sequence 812 BP; 188 A; 270 C; 209 G; 131 T; 14 other;

Query Match 26.9%; Score 746.4; DB 20; Length 812;
Best Local Similarity 97.2%; Pred. No. 2.5e-157;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;

QY 1570 CCAAGCCAGAGGAGCCCGCCGCCCAAGAGCCAGCAGCTGGACAGGGGCTGGGACCG 1629
Db 1 ccaagccagagaggcccccgccccaaggaagccagcactggacaaggggcctgggacccg 60
QY 1630 GCCAGGGTGCAGTGTTCAGGACCCCGCCGAAAGCTCTCTGGGACTGTGCTGGCACAGCCC 1689
Db 61 gccaggcgagtgctcaggaacccccccggaggcctcctggagctgctgctggaacagccc 120
QY 1690 GAGGCGCTGAAGTGGGAGCAGCAGGCTCAGTGCAGCACCCGAGCATACACCGCGCG 1749
Db 121 gagggcctgaagtgggcagcagcggtcaggtgcagcaccgcgagcatcacccaccgccc 180
QY 1750 AGGGTCCAGTGTCTACTTCCAGAGTGAAGATGAAGGCATGAAGGAGCTGCTGGTGG 1809
Db 181 aggggtccagtgtcactttccagagtgaagatgaaggagctggaaggagctgctgggtg 240
QY 1810 CCACCAAGATCAACTCGAGCGCCATCAAGCTCAACTCAAGGCACAGCTCCCAAGTGCAGA 1869
Db 241 ccaccaagatcaactcgagcgccatcaagctgcaactcaccgagcagctgcgaagtgcaga 300
QY 1870 TGAAGAGCAAGAAAGTGTCCACCCCTAGTACTACACTGTGTCTTTCCTCAAGCGGACG 1929
Db 301 tgaagaagcagaagaaagtgctccacccctagtagtacactctgtcttctcctcaagcgagc 360
```


10-APR-1997; 97US-0835913.
(GEMY) GENETICS INST INC.
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racie LA, Spaulding V, Treacy M;
WPI; 1999-070076/06.
New polynucleotides encoding human secreted proteins - derived from
e.g. human blood, kidney, foetal lung, placenta, testes, brain,
ovary, pituitary, retina and colon cDNA libraries
Claim 1; Page 291; 633pp; English.
This sequence represents an expressed sequence tag (EST), and is a
polynucleotide of the invention. The polynucleotides of the invention are
all secreted EST sequences isolated from a variety of human tissue
sources. The EST sequences and proteins encoded by them are predicted to
have useful biological activities which would make them suitable for
treating, preventing or ameliorating medical conditions in humans and
animals, although no supporting data is given. Suggested activities
include nutritional activity, immune stimulating or suppressing activity,
haematopoiesis regulating activity, tissue growth activity,
activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, receptor/ligand activity, anti-inflammatory
activity, cadherin/tumour invasion suppressor activity, tumour inhibition
activity. The EST sequences are also stated to be useful for gene
therapy.
Sequence 253 BP; 31 A; 106 C; 57 G; 59 T; 0 other;

Query Match	8.4%	Score 234	DB 20	Length 253
Best Local Similarity	100.0%	Pred. No. 3.4e-43		
Matches 234	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy 202	AGCCGGACCTGGTGTCTCTTTCTTTGGTGGGCTTCGTGGAGGACATTTTCTGGCTGTCA	261		
Db 20	agcggacctggtctccttctctgtgtctggtccttcgtggagcatcttctggtgtgtca	79		
Qy 262	ACCGGGTCATCCCTTACCACACTTCCGAGCTACACTTCCAGCCGACCCCGCCCGCACC	321		
Db 80	acccggtcatccttaccacagttcccgagctcaccttccagccagcccgcccccgacc	139		
Qy 322	CGCTGGCGGGCTCACCTACTTTCCCGTGGCGGACCTGTCTATATATCGCCGCCCTCTATG	381		
Db 140	cgctggcggtcacctacttccgttgccgacctgtctatcatcgccgccctctatg	199		
Qy 382	CCCGGTTACCGCCAGATCGAGCGCGGCTGCACCTGTCCCTCTATCCTCGAG	435		
Db 200	cccgcttcacgcgcagatccgagcgcgtcgacctgtccctctatcctcag	253		

RESULT	8	
AAT22851		
ID	AAT22851	standard; cDNA to mRNA; 67 BP.
XX	XX	
XX	AC	AAT22851;
XX	XX	
DT	29-AUG-1996	(first entry)
XX	XX	
DE	Human	gene signature HUMGS04527.
XX	XX	
KW	Gene	signature; messenger RNA; mRNA; relative abundance; frequency;
KW	human;	cloning; mapping; non-biased library; diagnosis; detection;
XX	cell	typing; abnormal cell function; ss.
XX	XX	
OS	Homo	sapiens.
XX	XX	
PN	WO9514772-A1.	
XX	XX	

PD	01-JUN-1995.	
XX		
XX	11-NOV-1994;	94WO-JP01916.
XX		
XX	12-NOV-1993;	93JP-0355504.
XX		
PA	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
XX		
XX	Matsubara K, Okubo K;	
PI		
PI	WPT: 1995-206931/27.	
DR		
XX		
XX	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	
PT		
XX		
XX	Claim 1; Page 1227; 2245pp; Japanese.	
PS		
XX		
CC	A single-stranded DNA (or its complementary strand or the corresp.	
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences	
CC	given in ARI9001-r26837 and which is able to hybridise to part of	
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)	
CC	sequences were obtained from 3'-directed cDNA libraries prepared	
CC	from various human tissues; synthesis of cDNA was initiated from the	
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-	
CC	untranslated sequence is unique to a particular mRNA species, almost	
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library	
CC	is constructed so as to reflect accurately the relative abundance of	
CC	different mRNAs in the particular tissue from which it was derived.	
CC	The appearance frequency of a given GS in a cDNA library can be	
CC	determined (esp. using primers and probes derived from the GS	
CC	sequences) as a means of diagnosing abnormal cell function or for	
CC	recognising different cell types.	
XX		
XX	Sequence 67 BP: 20 A; 9 C; 9 G; 28 T; 1 other.	
SQ		

	Query Match	2.48;	Score 66;	DB 16;	Length 67;
	Best Local Similarity	98.5%;	Pred. No. 6.7e-06;		
	Matches 66;	Conservative	0; Mismatches 1;	Indels 0;	Gaps 0;
Qy	2693	GATCTCTCCGTGTTTGTACATAGATTATTTTTTCAGTTCACAAGAAAGATCAATACATTT	2752		
Dd	1	gattcttccgttttgtacatagatttattttcagtgtccaagaagaatgaatacattt	60		
Qy	2753	TGTTAAA	2759		
Dd	61	tgtnaaa	67		

RESULT	9	
AAZ87297		
ID	AAZ87297	standard; DNA; 13842 BP.
XX		
XX	AAZ87297;	
XX		
DT	05-JUN-2000	(first entry)
XX		
DE	S. venezuelae	macrolide biosynthetic gene pikAI, SEQ ID NO:30.
XX		
KW	Desosamine	biosynthesis; macrolide; polyketide; methymycin; pikromycin;
KW	neomethymycin;	narbomycin; polyhydroxyalkanoate monomer synthase;
KW	biopolymer;	antibiotic; chemotherapeutic; immunosuppressant; asthma,
KW	chronic obstructive	pulmonary disease; respiratory inflammation;
KW	hypercholesterolaemia;	crop protection agent; ds.
XX		
OS	Streptomyces	venezuelae ATCC15439.
XX		
Key	Location/Qualifiers	
FH	1..13842	
CDS	/*tag= a	
FT		
FT		

CC monomers. The compounds produced by the recombinant host cells are useful
CC as biopolymers, e.g., in packaging or biomedical applications, to
CC engineer PBA monomer syntheses or to prepare biologically active agents,
CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
CC chronic obstructive pulmonary disease as well as other diseases involving
CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
CC antibiotics which are active against a variety of organisms, e.g.,
CC bacteria, including multi-drug resistant pneumococci and other
CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
CC protection agents (e.g., fungicides or insecticides) via expression of
CC polyketides in plants. The present sequence represents the macrolide
CC biosynthetic gene cluster (pik) from *Streptomyces venezuelae* ATCC 15439.
XX
SQ Sequence 37948 BP: 4901 A; 14718 C; 13311 G; 5018 T; 0 other;

[illegible]

RESULT	12
AAA75633	
ID	AAA75633 standard; DNA; 38506 BP.
XX	
AC	AAA75633;
XX	
DT	22-JAN-2001 (first entry)

DE	Nucleotide sequence of the insert DNA in cosmid pKOS023-27.
DX	
XX	
KW	Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;
KW	antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;
KW	desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
KW	picromycin biosynthesis; ss.
XX	
OS	Streptomyces venezuelae.
PN	US6117659-A.
XX	
PD	12-SEP-2000.
XX	
XX	
PF	27-MAY-1999; 99US-0320878.
XX	
XX	
PR	28-MAY-1998; 98US-0087080.
PR	22-SEP-1998; 98US-0100880.
PR	08-FEB-1999; 99US-0119139.
PR	20-MAY-1999; 99US-0134990.
PR	30-APR-1997; 97US-0846247.
PR	06-MAY-1998; 98US-0073538.
PR	28-AUG-1998; 98US-0141908.
XX	
XX	(KOSA-) KOSAN BIOSCIENCES INC.
PA	
XX	
PI	Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;
XX	
DR	WPI; 2000-610844/58.
XX	
PT	New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
PT	for converting ketolides to antibiotics and as antibiotics and
PT	intermediates in the synthesis of compounds with pharmaceutical value
PT	
XX	
XX	Disclosure; Columns 15-32; 117pp; English.
PS	
XX	
CC	The present sequence is used to produce the recombinant DNA compounds
CC	of the invention. The specification describes a recombinant DNA compound
CC	expressing recombinant polyketide synthase genes in host cells for the
CC	production of narbonolide, narbonolide derivatives and polyketides that
CC	are useful as antibiotics and as intermediates in the synthesis of
CC	compounds with pharmaceutical value. The DNA compounds may also encode
CC	a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
CC	transferase enzymes (useful for conversion of ketolides to antibiotics)
CC	and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
CC	These compounds are also useful for increasing the antibiotic activity
CC	of a compound relative to the unhydroxylated compound. The recombinant
CC	host cells are useful as genetic systems that allow rapid engineering
CC	of the narbonolide polyketide synthase. These would be valuable for
CC	creating novel ketolide analogs for pharmaceutical applications.
XX	
SO	Sequence 38506 BP:4914 A; 15118 C; 13444 G; 5030 T; 0 other;

	Query Match	1.98;	Score 52.6;	DB 21;	Length 38506;
	Best Local Similarity	45.3%;	Pred. No. 0.052;		
	Matches 274;	Conservative 0;	Mismatches 324;	Indels 7;	Gaps 2;
Qy	52	CGGGCCGCGCACTATTTCCAGGCTCTGCGGGGCAGGGGCCGCGCC-CACCGCCGCGCCGCC	110		
Db	9676	cggccgcccgcactgtctcgtg caactcggttcggcgagatcgcgcgcgcacgtcgcgcgtt	9735		
Qy	111	ATGGGGCTGAAGCGCGCCAGAAAGCAGTGTCCTCCGCTCATYCGACAGCTGGTGTG	170		
Db	9736	gtgttctcgtcgcgcagcgcgcgccttgctgcgcgcgcgcgcgcgtcatgcaggag	9795		
Qy	171	CGCCTGTTTGTCGCGAGCTCGGGCGAGAGAGCGCGACCTGGTGCTCTCTTTCTTGCTG	230		
Db	9796	ctgccgcgcggctggcgcgatgctctcgcctccaggccgcggaggacagagatccgcgctgtg	9855		
Qy	231	CTGGGCTTCGGGAGACATTTTCTGGCTGTCAACCGCGTCATCCCCTACCACAGTTCCCGAG	290		
Db	9856	ctgaagaacgaagaaacqatacqcaqaacqctctgaacqctcgcgcgcgtctaaacggccccag	9915		

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 19:53:04 ; Search time 15359.2 seconds
(without alignments)
2791.583 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTCGCGGAGCGCGGACC.....TGTAAAAA..... 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_ba1:*
17: gb_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_ba1:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_vl3:*
61: gb_vl4:*
62: gb_vl5:*
63: gb_vl6:*
64: gb_vl7:*
65: gb_vl8:*
66: gb_vl9:*
67: gb_vl10:*
68: gb_vl11:*
69: gb_vl12:*
70: gb_vl13:*
71: gb_vl14:*
72: gb_vl15:*
73: gb_vl16:*
74: gb_vl17:*
75: gb_vl18:*
76: gb_vl19:*
77: gb_vl20:*
78: gb_vl21:*
79: gb_vl22:*
80: gb_vl23:*
81: gb_vl24:*
82: gb_vl25:*
83: gb_vl26:*
84: gb_vl27:*
85: gb_vl28:*
86: gb_vl29:*
87: gb_vl30:*
88: gb_vl31:*
89: gb_vl32:*
90: gb_vl33:*
91: gb_vl34:*
92: gb_vl35:*
93: gb_vl36:*
94: gb_vl37:*
95: gb_vl38:*
96: gb_vl39:*
97: gb_vl40:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2772	100.0	2772	97	HSU93236
2	2406	86.8	2633	91	BC002544
3	1865	2	67.3	9	AR081879
4	1781	64.2	1837	97	HSURF
5	1648	6	59.5	3067	94
6	1622	58.5	1691	91	BC002664
7	1616	4	58.3	2984	94
8	1616	58.3	2623	94	AF016398

	577 a	909 c	779 g	507
BASE COUNT				
ORIGIN				
exon	/number=6 1023. .1159 /gene="MEN1" /number=7 1132 /gene="MEN1" /standard_name="l132delG" /note="1 bp deletion" /replace="" 1160. .1295 /gene="MEN1" /number=8 1197. .1199 /gene="MEN1" /standard_name="E363del" /note="3 bp deletion" /replace="" 1296. .1460 /gene="MEN1" /number=9 1416 /gene="MEN1" /standard_name="W436R" /note="Trp436Arg" /replace="C" 1417 /gene="MEN1" /standard_name="W436X" /note="Trp436stop" /replace="A" 1461. .2772 /gene="MEN1" /number=10 1689 /gene="MEN1" /standard_name="R527X" /note="Arg527stop" /replace="T"			
variation				
exon				
variation				
exon				
variation				
exon				
variation				

Query Match	100.0%;	Score 2772;	DB 97;	Length 2772;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2772;	Conservative	0;	Mismatches	0;
			Indels	Gaps
Qy	1	GGTCTCGGAGCCGCGACCTAGAGATCCACAGAGCCACAGCGACGCGCCGCGCCGCC	60	
Db	1	GGTCTCGGAGCCGCGGACCTTAGAGATCCACAGAGCCACAGCGACGCGCCGCGCCGCC	60	
Qy	61	ACTATTTCCAGGCTCTGCGGGCGAGGGCCGCCGCCACCGCCGCCGCCCATGGGCGCTGA	120	
Db	61	ACTATTTCCAGGCTCTGCGGGCGAGGGCCGCCGCCACCGCCGCCGCCCATGGGCGCTGA	120	
Qy	121	AGGCGGCCACAGACGCTGTTCGCTGCGCTCCATCGAGACGCTGTCGCCCTGTTTG	180	
Db	121	AGGCGGCCACAGAACGCTGTTCGCGCTGCGCTCCATCGAGACGCTGTCGCCCTGTTTG	180	
Qy	181	CTGCCGAGCTGGGCGCGAGAGAGCGGACCTGGTGCTCCTTTCCCTTGCTGGGCGCTCG	240	
Db	181	CTGCCGAGCTGGGCGCGAGAGAGCGGACCTGGTGCTCCTTTCCCTTGCTGGGCGCTCG	240	
Qy	241	TGGAGCATTTTCTGGCTGTCAACCGCGTCAATCCCTACCAAGCTTCCCGAGCTCACCTCC	300	
Db	241	TGGAGCATTTTCTGGCTGTCAACCGCGTCAATCCCTACCAAGCTTCCCGAGCTCACCTCC	300	
Qy	301	AGCCAGCCCGCCGCCGCCGACCGCTGGCGGCTCCACCTACTTTCCCGTGGCGACCTGT	360	
Db	301	AGCCAGCCCGCCGCCGCCGACCGCTGGCGGCTCCACCTACTTTCCCGTGGCGACCTGT	360	
Qy	361	CTATCATCGCGCCCTCTATCCCGCTTCAACCGCCAGATCCGAGGCGCCGTCGACCTGT	420	
Db	361	CTATCATCGCGCCCTCTATCCCGCTTCAACCGCCAGATCCGAGGCGCCGTCGACCTGT	420	

Qy	421	CCCTCTATCTCGAGAAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAGGTCTTCGATG	480
Db	421	CCCTCTATCTCGAGAAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAGGTCTCCGATG	480
Qy	481	TCATATGGAAACGCCTCAGCCGCTCCTACTTCAAGSATCGGCGCCACATCCAGTCCTCT	540
Db	481	TCATATGGAAACGCCTCAGCCGCTCCTACTTCAAGSATCGGCGCCACATCCAGTCCTCT	540
Qy	541	TCAGCTTTCATCAGAGCACCAAAATTGGACAGCTCCGGTGTGGCCCTTGTCTGTGTGGG	600
Db	541	TCAGCTTTCATCAGAGCACCAAAATTGGACAGCTCCGGTGTGGCCCTTGTCTGTGTGGG	600
Qy	601	CTTGCCAGGCGCTGGGTCTCCGGGATGTCACCTCGCCCTGCTTGAGGATCATGCCCTGG	660
Db	601	CTTGCCAGGCGCTGGGTCTCCGGGATGTCACCTCGCCCTGCTTGAGGATCATGCCCTGG	660
Qy	661	TAGTGTTTGGGCCCCAATGGGGAGCAGACAGCTGAGGTCACTGGCAGCGGAAGGCAACG	720
Db	661	TAGTGTTTGGGCCCCAATGGGGAGCAGACAGCTGAGGTCACTGGCAGCGGAAGGCAACG	720
Qy	721	AGGACGCGAGGGGCCAGACAGTCAATGCGGTGTGGCTGAGCGGAGTGGCTGTACCTGA	780
Db	721	AGGACGCGAGGGGCCAGACAGTCAATGCGGTGTGGCTGAGCGGAGTGGCTGTACCTGA	780
Qy	781	AAGGATCATACATCGCTGTGACCGCAAGATGGAGTGGCGTTCATGTTGTGCCATCA	840
Db	781	AAGGATCATACATCGCTGTGACCGCAAGATGGAGTGGCGTTCATGTTGTGCCATCA	840
Qy	841	ACCTTCCATTGACCTGTCACACCGACTCGCTGGAGCTTTCGAGCTGCAGCAAGCTGC	900
Db	841	ACCTTCCATTGACCTGTCACACCGACTCGCTGGAGCTTTCGAGCTGCAGCAAGCTGC	900
Qy	901	TCGTGGCTGCTATGACCTTGGACATCTGTGAAAGTATCCCATGGCTTAGGGAACCTGG	960
Db	901	TCGTGGCTGCTATGACCTTGGACATCTGTGAAAGTATCCCATGGCTTAGGGAACCTGG	960
Qy	961	CAGATCTAGAGGAGCTGAGGCCACCCCTGGCGGCCAGACCCACTCACCTCTACCACA	1020
Db	961	CAGATCTAGAGGAGCTGAGGCCACCCCTGGCGGCCAGACCCACTCACCTCTACCACA	1020
Qy	1021	AGGGCATGTGCTCAGCCAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACC	1080
Db	1021	AGGGCATGTGCTCAGCCAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACC	1080
Qy	1081	TGGCTGGCTACCACTGTCCCAACCGCAATGTGCGGGAAGCCCTGCAGCCTTGGCGGACA	1140
Db	1081	TGGCTGGCTACCACTGTCCCAACCGCAATGTGCGGGAAGCCCTGCAGCCTTGGCGGACA	1140
Qy	1141	CGGCCACTGTCTATCCAGGACTACAACCTACTGCGGGGAAGACGAGAGATCTACAAGGAT	1200
Db	1141	CGGCCACTGTCTATCCAGGACTACAACCTACTGCGGGGAAGACGAGAGATCTACAAGGAT	1200
Qy	1201	TCCTTGAAGTAGCAATGATGTCTCCCACTGCTCAAGAGGCGACGAGCTTGTCTGG	1260
Db	1201	TCCTTGAAGTAGCAATGATGTCTCCCACTGCTCAAGAGGCGACGAGCTTGTCTGG	1260
Qy	1261	AGCGGCGGAGGAGCGCGGGGGAGCAAGCCAGGGCACCCAGAGCCAGGTTCCGCC	1320
Db	1261	AGCGGCGGAGGAGCGCGGGGGAGCAAGCCAGGGCACCCAGAGCCAGGTTCCGCC	1320
Qy	1321	TCCAGGACCTGTAGTGTCTTGGCCCACTGCTGCGATTCTACGACGGCATCTGCAAAATGG	1380
Db	1321	TCCAGGACCTGTAGTGTCTTGGCCCACTGCTGCGATTCTACGACGGCATCTGCAAAATGG	1380
Qy	1381	AGGAGGACGTCCACCGCTGTGTGTCAGCTGGGCTGGGCCACCTTTCTGTGTGAGTCCC	1440
Db	1381	AGGAGGACGTCCACCGCTGTGTGTCAGCTGGGCTGGGCCACCTTTCTGTGTGAGTCCC	1440
Qy	1441	TAGCCGCTTTTGTAGGACAGAGTGTGCGCGAGAAGGTGCGCATAGTGAGCGCGAGGCCGAGG	1500
Db	1441	TAGCCGCTTTTGTAGGACAGAGTGTGCGCGAGAAGGTGCGCATAGTGAGCGCGAGGCCGAGG	1500

EHFLAVNRVITPNVPELTQPPAPPPGGTLTYPVADLSIAALYARFQAIRGADV
VSLYPRREGVSRLVKVSDVWINSLSRSYFKDRAHQISLFSFTTGTGKLSSGVAPA
LVGCAOLGLRDVHLALSDHAWKLYLKGYSMRCDKMEVAEMVAINPISDIDHTDS
LELLQOKLLWLLDGLHLERYPMALGNLADLELEPTGRPDPLTLYHKGIASAKT
YREBHIYPYMLAGYHCRNRNRREALQAWADATV IODYNYCREDEITYKEFFEVA
DYINLUKEAASLLLEAGERPEQSQSALQDPECFALHLLRFVDDGICKWEEBS
PTVLHVGMATFLVLSGRFRFGQVRKYVIRSEAEAAEPEPMGEAREGRGRPR
ESKPEPPPKPLDGLGTGOGAVSPRPPTVAGTARGPEGSGTAQVPAPAS
PPPEGPVLTFSQSEKMKMKELLVATKINSSAIKLQTLAQSQVQMKKQKRVSPSYTSL
FLKQRKGL"

BASE COUNT 548 a 865 c 729 g 491 t
ORIGIN

Query Match 86.8%; Score 2406; DB 91; Length 2633;
Best Local Similarity 95.7%; Pred No. 0;
Matches 2579; Conservative 0; Mismatches 5; Indels 110; Gaps 6;

Qy	80	GGCGAGGCGCGCCGCCACCCGCCCATGGGCTGAAGGCGCGCCAGAACGCGT	139
Db	47	GGCGAGGCGCGCGCCGCCACCCGCCCATGGGCTGAAGGCGCGCCAGAACGCGT	106
Qy	140	GTTCGCGCTGCCTCATCAGCAGCGTGTGCGCCTGTTTGTGCGGAGCTGGGCGGAGA	199
Db	107	GTTCGCGCTGCCTCATCAGCAGCGTGTGCGCCTGTTTGTGCGGAGCTGGGCGGAGA	166
Qy	200	GGAGCGGACCTGGTGCTCTTCTTGTGCTGGGCTTCGTGGAGCATTTTCTGCGTGT	259
Db	167	GGAGCGGACCTGGTGCTCTTCTTGTGCTGGGCTTCGTGGAGCATTTTCTGCGTGT	226
Qy	260	CAACGCGCTCATCCTTACCAAGCTTCCCGAGCTCACTTCCAGCCAGCCCGCCCGCA	319
Db	227	CAACGCGCTCATCCTTACCAAGCTTCCCGAGCTCACTTCCAGCCAGCCCGCCCGCA	286
Qy	320	CCGCGCTGCGCGCTCACTTCCCGTGGCGGACCTGTCTATCATCGCGCGCTCTA	379
Db	287	CCGCGCTGCGCGCTCACTTCCCGTGGCGGACCTGTCTATCATCGCGCGCTCTA	346
Qy	380	TGCGCGCTTCAACGCGCCAGATCCGAGGCGCGTCCAGCTTCTTCTTATCTCGAGAAG	439
Db	347	TGCGCGCTTCAACGCGCCAGATCCGAGGCGCGTCCAGCTTCTTCTTATCTCGAGAAG	406
Qy	440	GGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCAATGAAACAGCCTCAG	499
Db	407	GGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCAATGAAACAGCCTCAG	466
Qy	500	CGCTCTCTTCAAGGATCGGCGCCACATCCAGTCCCTTTCAGCTTCTACAGGCGAC	559
Db	467	CGCTCTCTTCAAGGATCGGCGCCACATCCAGTCCCTTTCAGCTTCTACAGGCGAC	526
Qy	560	CAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCGCTGCCAGGCGCTGGTCT	619
Db	527	CAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCGCTGCCAGGCGCTGGTCT	586
Qy	620	CGGGATGTCCACCTCGCCCTGTGTAGGATCATGCTGGGTAGTGTGGGCCCAATGG	679
Db	587	CGGGATGTCCACCTCGCCCTGTGTAGGATCATGCTGGGTAGTGTGGGCCCAATGG	624
Qy	680	GGAGCAGACAGCTGAGGTACCTTGGCGACCGCAAGGCGACGAGGCGCGAGGCGCGAC	739
Db	625	-----	624
Qy	740	AGTCAATGCCGTGTGGCTGAGCGGAGCTGGCTGTACCTGAAAGGATCATACATGCGCTG	799
Db	625	-----GGAGCTGGCTGTACCTGAAAGGATCATACATGCGCTG	661
Qy	800	TGACCCCAAGATGGAGGTGGCGTTATGTGTGTGCGCATCAACCCCTTCCATGACCTGCA	859
Db	662	TGACCCCAAGATGGAGGTGGCGTTATGTGTGTGCGCATCAACCCCTTCCATGACCTGCA	721
Qy	860	CACCGACTCGCTGGAGCTTCTCAGCTCAGCAGAGCTGCTTGTGCTCTATGACCT	919
Db	722	CACCGACTCGCTGGAGCTTCTCAGCTCAGCAGAGCTGCTTGTGCTCTATGACCT	781

Qy	920	GGGACATCTGGAAGGTACCCCATGSGCCTTAGGAACTGGCAGATCTAGAGGAGCTGCA	979
Db	782	GGGACATCTGGAAGGTACCCCATGSGCCTTAGGAACTGGCAGATCTAGAGGAGCTGCA	841
Qy	980	GCCCAACCCCTGCGCGGCCAGACCCACTCACCTCTACCAACAAGGGCATTTGCCCTAGCCAA	1039
Db	842	GCCCAACCCCTGCGCGGCCAGACCCACTCACCTCTACCAACAAGGGCATTTGCCCTAGCCAA	901
Qy	1040	GACCTACTACTGGGATGAACACATCTACCCCTTACATGTACCTGGCTGGCTTACCACTGTGC	1099
Db	902	GACCTACTACTGGGATGAACACATCTACCCCTTACATGTACCTGGCTGGCTTACCACTGTGC	961
Qy	1100	CAACGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGGCCACTGTTCATCCAGGA	1159
Db	962	CAACGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGGCCACTGTTCATCCAGGA	1021
Qy	1160	CTACAACCTACTGCCGGGAAGAGAGGATCTACAAGAGTCTTTTGAAGTAGCCAAATGA	1219
Db	1022	CTACAACCTACTGCCGGGAAGAGAGGATCTACAAGAGTCTTTTGAAGTAGCCAAATGA	1081
Qy	1220	TGTATCCCAACCTGCTGAAGGAGCAGCAGCTTGTCTGGAGGGCGGCGAGGAGCGGC	1279
Db	1082	TGTATCCCAACCTGCTGAAGGAGCAGCAGCTTGTCTGGAGGGCGGCGAGGAGCGGC	1141
Qy	1280	GGGGAGGAGCAAGGCGCAGAGCCAGAGCCCAAGTTCGCCCTTCCAGGACCCCTGAGTGT	1339
Db	1142	GGGGAGGAGCAAGGCGCAGAGCCAGAGCCCAAGTTCGCCCTTCCAGGACCCCTGAGTGT	1201
Qy	1340	CGCCCACTTGTGCGATTTACGAGGCGATCTCAAAATGGAGGGGAGTCCCAAGCC	1399
Db	1202	CGCCCACTTGTGCGATTTACGAGTGGCATCTCAAAATGGAGGGGAGTCCCAAGCC	1261
Qy	1400	TGTGTGACAGTGGGCTGGGCGCACCTTTCTTGTGCACTCTCTAGGCGCTTTTGAGGGACA	1459
Db	1262	TGTGTGACAGTGGGCTGGGCGCACCTTTCTTGTGCACTCTCTAGGCGCTTTTGAGGGACA	1321
Qy	1460	GGTGGGCGAAGAGTGGCGATAGTGGCGAGAGGCGGCGGCGGCGGCGGCGGAGGCGC	1519
Db	1322	GGTGGGCGAAGAGTGGCGATAGTGGCGAGAGGCGGCGGCGGCGGCGGCGGAGGCGC	1381
Qy	1520	GTGGGCGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGAGTCCAAAGCAGA	1579
Db	1382	GTGGGCGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCGGCGGCGGAGTCCAAAGCAGA	1441
Qy	1580	GGAGCCCGCGCGCCCAAGAGCCAGCTGGACAGGCGCTGGGCGACCGCGCGGAGGTCG	1639
Db	1442	GGAGCCCGCGCGCCCAAGAGCCAGCTGGACAGGCGCTGGGCGACCGCGCGGAGGTCG	1501
Qy	1640	AGTGTGAGGACCCCGCGGAAGCTCTTGGGACTGTGCTGGGCACAGCCCGAGGCGCTGA	1699
Db	1502	AGTGTGAGGACCCCGCGGAAGCTCTTGGGACTGTGCTGGGCACAGCCCGAGGCGCTGA	1561
Qy	1700	AGGTGGCAGCAGGCTCAGGTGCCAGCAGCCCGCAGCATCACACCCCGGAGGCTCCAGT	1759
Db	1562	AGGTGGCAGCAGGCTCAGGTGCCAGCAGCCCGCAGCATCACACCCCGGAGGCTCCAGT	1621
Qy	1760	GCTCACTTTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGTGGCCACCAAGAT	1819
Db	1622	GCTCACTTTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGTGGCCACCAAGAT	1681
Qy	1820	CAACTCGAGCGCATCAAGCTGCACTACGGCAGCTGCGCAGTGCAGTGAAGAAGCA	1879
Db	1682	CAACTCGAGCGCATCAAGCTGCACTACGGCAGCTGCGCAGTGCAGTGAAGAAGCA	1741
Qy	1880	GAAAGTGTCCACCCCTAGTACTACTCTGTCTTTCTCAAGCGGCGAGCGGAAGGCT	1939
Db	1742	GAAAGTGTCCACCCCTAGTACTACTCTGTCTTTCTCAAGCGGCGAGCGGAAGGCT	1801
Qy	1940	CTGAACCTACTGGGAGCTTCGAGCCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCCA	1998
Db	1802	CTGAACCTACTGGGAGCTTCGAGCCGCTTGTGGGACCCAGGCTCCGCTTAGTCCCCA	1861

Qy	1999	ACTCTAGCCCATGTTCTGCCCCAGCCCAAGAGGGAGAGCCCTCACTCTACCCAAACC	2058
Db	1862	ACTCTAGCCCATGTTCTGCCCCAGCCCAAGAGGGAGAGCCCTCACTCTACCCAAACC	1921
Qy	2059	CTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAACCCACGATTTTCTCAGCTCAGAAC	2118
Db	1922	CTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAACCCACGATTTTCTCAGCTCAGAAC	1981
Qy	2119	CCAGGGCTCTGCCCGAGTCTGTAGATATAGGTCTCTTCTCCAGAGATCCAGCCGGCCA	2178
Db	1982	CCAGGGCTCTGCCCGAGTCTGTAGATATAGGTCTCTTCTCCAGAGATCCAGCCGGCCA	2041
Qy	2179	ATGGAACCTCAGCTGGGTCTTAATACCAGTCTTTAAAGGCCAGCCCTCAGAAACC	2238
Db	2042	ATGGAACCTCAGCTGGGTCTTAATACCAGTCTTTAAAGGCCAGCCCTCAGAAACC	2101
Qy	2239	AGCTCTCTCTCGGAACCGCTCACTAGAGCCAGACCAAGTTTACTCAGGGCTCTCTCCA	2298
Db	2102	AGCTCTCTCTCGGAACCGCTCACTAGAGCCAGACCAAGTTTACTCAGGGCTCTCTCCA	2161
Qy	2299	GCTTGTAGGAGCTGAGGTTTACCCCTTTAACCCAAAGGAGGACACAGGTCCCACTCCAGCCC	2358
Db	2162	GCTTGTAGGAGCTGAGGTTTACCCCTTTAACCCAAAGGAGGACACAGGTCCCACTCCAGCCC	2220
Qy	2359	GGGAGGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAAATTCATATC	2418
Db	2221	GGGAGGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAAATTCATATC	2279
Qy	2419	TTGGAATCAAGTCTCCTGCCCAATAACTTCAGTCTCTTCTCCAGAAATTTGGAATC	2478
Db	2280	TTGGAATCAAGTCTCCTGCCCAATAACTTCAGTCTCTTCTCCAGAAATTTGGAATC	2338
Qy	2479	CTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAACCTCCGCCCGCCAGCTATGA	2538
Db	2339	CTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAACCTCCGCCCGCCAGCTATGA	2398
Qy	2539	GCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCGCCGATCAGAGAGACCTCCC	2598
Db	2399	GCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCGCCGATCAGAGAGACCTCCC	2458
Qy	2599	TTCCGACCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGAGCATCCCGAGGAAT	2658
Db	2459	TTCCGACCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGAGCATCCCGAGGAAT	2518
Qy	2659	CTGACAGGGGTTAGGAGTGGGTGACAGAGCTGATCTCTCTCTGTTTGTACATAGAT	2718
Db	2519	CTGACAGGGGTTAGGAGTGGGTGACAGAGCTGATCTCTCTCTGTTTGTACATAGAT	2577
Qy	2719	TTATTTTTCAGTCCCAAGAAAGATGAATACATTTTGTAAAAAATAAAAAA	2772
Db	2578	TTATTTTTCAGTCCCAAGAAAGATGAATACATTTTGTAAAAAATAAAAAA	2631
RESULT 3			
AR081879			
LOCUS	2196 bp	DNA	PAT
DEFINITION	Sequence 2 from patent US 5972649.		31-AUG-2000
ACCESSION	AR081879		
VERSION	AR081879.1	GI:10008605	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2196)		
AUTHORS	Au-Young, J., Covitz, P.A., Tang, Y. Tom and Murrey, L.E.		
TITLE	Polynucleotide sequence encoding human multiple endocrine neoplasia type 1 protein		
JOURNAL	Patent: US 5972649-A 2 26-OCT-1999;		
FEATURES	Location/Qualifiers		
source	1. 2196		
BASE COUNT	445 a 712 c 623 g 385 t		31 others
ORIGIN			

Query Match 67.3%; Score 1865.2; DB 9; Length 2196;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2082; Conservative 7; Mismatches 50; Indels 140; Gaps 6;

Qy	88	GCAGCGCCACCGCCGCGCCATGGGGTGAAGGCCGCCAGAGACGCTTTTCCCGC	147
Db	35	GCAGCGCCACCGCCGCGCCATGGGGTGAAGGCCGCCAGAGACGCTTTTCCCGC	94
Qy	148	TGCGCTCCATCGAGGAGCTGGTGGCTGTTTGTGCCGAGCTGGGCCGAGAGGACCGG	207
Db	95	TGCGCTCCATCGAGGAGCTGGTGGCTGTTTGTGCCGAGCTGGGCCGAGAGGACCGG	154
Qy	208	ACCTGGTGTCTCTTCTTCTGGCTTCGTGGAGCATTTTCTGGCTCTCAACCGCG	267
Db	155	ACCTGGTGTCTCTTCTTCTGGCTTCGTGGAGCATTTTCTGGCTCTCAACCGCG	214
Qy	268	TCATCCCTACCAAGTTCCGAGCTCACCTTCCAGGCCAGCCGCCGCCAGCCGGCTG	327
Db	215	TCATCCCTACCAAGTTCCGAGCTCACCTTCCAGGCCAGCCGCCGCCAGCCGGCTG	274
Qy	328	GGCGCTCACTACTTTCCTGGCTTCGTGGAGCATTTTCTGGCTCTCAACCGCG	387
Db	275	GGCGCTCACTACTTTCCTGGCTTCGTGGAGCATTTTCTGGCTCTCAACCGCG	334
Qy	388	TCACCGCCAGATCCGAGCGCGCTGACCTGTCCCTCTATCTCGAGAGGGGTGTCT	447
Db	335	TCACCGCCAGATCCGAGCGCGCTGACCTGTCCCTCTATCTCGAGAGGGGTGTCT	394
Qy	448	CCAGCGTGTAGCTGGTGAAGAGTCTCCGATGTATATGGAACAGCTTCAGCGCTCT	507
Db	395	CCAGCGTGTAGCTGGTGAAGAGTCTCCGATGTATATGGAACAGCTTCAGCGCTCT	454
Qy	508	ACTTCAGGATCGGGCCACATCCAGTCCCTCTTCAGCTTCATCACAGGACCAATGG	567
Db	455	ACTTCAGGATCGGGCCACATCCAGTCCCTCTTCAGCTTCATCACAGGACCAATGG	514
Qy	568	ACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCTGCCAGGCTCGGCTCTCGGGATG	627
Db	515	A---CTCCGNGTGGCTTTGCT-----	534
Qy	628	TCCACCTCGCCCTGTCTGAGGATCATGCCGTGGGTAGTGTGGGCCCAATGGGGAGCAGA	687
Db	535	-----GTGNNTGGGCCCAATGGGGAGCAGA	559
Qy	688	CAGCTGAGGTACCTGGCAGCGCAAGGGCAACGAGGACCGAGGGCCAGACAGTCAATG	747
Db	560	CAGCTGAGGTACCTGGCAGCGCAAGGGCAACGAGGACCGAGGGCCAGACAGTCAATG	619
Qy	748	CCGCTGTGGCTGAGCGGAGCTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA	807
Db	620	CCGCTGTGGCTGAGCGGAGCTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA	679
Qy	808	AGATGAGGTGGGCTTTCATGGTGTGCCATCAACCCCTTCCATTTGAGCTGCACACCACT	867
Db	680	AGATGAGGTGGGCTTTCATGGTGTGCCATCAACCCCTTCCATTTGAGCTGCACACCACT	739
Qy	868	CGTGTGAGCTTCTCCAGCTGCAGCAGAGCTGCTGCTGCTCTATGCTGAGCTGGGATC	927
Db	740	CGTGTGAGCTTCTCCAGCTGCAGCAGAGCTGCTGCTGCTCTATGCTGAGCTGGGATC	799
Qy	928	TGGAAGGTATACCCCATGGCTTTAGGGAACCTTGGCAGATCTAGAGGAGCTGCAGCCACCC	987
Db	800	TGGAAGGTATACCCCATGGCTTTAGGGAACCTTGGCAGATCTAGAGGAGCTGCAGCCACCC	859
Qy	988	CTGCGCCGCGCAGACCCACTCACCTCTACACAAAGGCGATTCCTCAGCCAGACCTACT	1047
Db	860	CTGCGCCGCGCAGACCCSRGGCGCTGCTRCACAAAGGCGATTTGCTTTCAGCCAAAGACCTACT	919
Qy	1048	ATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCTGTGCGCAACCGCA	1107
Db	920	ATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTGTGCGCAACCGCA	979

```
QY 1108 ATGTGGGGAAGCCCTGACGCTCTGGG---CGGACACGGCCACTGTCTACCCAGGACTACA 1164
Db 980 ATGTGGGGAAGCCCTGACGCTCTGGGACAGACACAGGCACTCTCTCCAGGACTACA 1039
QY 1165 ACTACTGCGGGGAAGCAGGAGATCTCAAGAGATTTCTTTGAAGTAGCAATGATGCA 1224
Db 1040 ACTACTGCGGGGAAGCAGGAGATCTCAAGAGATTTCTTTGAAGTAGCAATGATGCA 1099
QY 1225 TCCCAACCTGCTGAAGAGGACGACGCTTCTGAGAGCGGCGGAGGAGCGCGGGG 1284
Db 1100 TCCCAACCTGCTGAAGAGGACGACGCTTCTGAGAGCGGCGGAGGAGCGCGGGG 1159
QY 1285 AGCAAGAGCGGACCCAGAGCCCAAGTTCTCGCCCTCCAGGACCTGAGTCTTCGCCC 1344
Db 1160 AGCAAGAGCGGACCCAGAGCCCAAGTTCTCGCCCTCCAGGACCTGAGTCTTCGCCC 1219
QY 1345 ACCTGCTGGATTCTACGAGCGCATCTGCAATGGAGAGGAGGAGTCCAGCGCTGTGC 1404
Db 1220 ACCTGCTGGATTCTACGAGCGCATCTGCAATGGAGAGGAGGAGTCCAGCGCTGTGC 1279
QY 1405 TGCAGTGGGCTGGCCACCTTTCTTGTGCTGCTAGGCGGTTTGGAGGACAGGTGC 1464
Db 1280 TGCAGTGGGCTGGCCACCTTTCTTGTGCTGCTAGGCGGTTTGGAGGACAGGTGC 1339
QY 1465 GGCAGAAGGTGGCAGATAGTGGAGGAGGCGGCGGCGGCGGAGGCGGAGGCGGTGG 1524
Db 1340 NSCAGAAGGTGGCAGATAGTGGAGGAGGCGGCGGCGGAGGCGGAGGCGGTGG 1524
QY 1525 GCGAGGAAGCGGGAAGCGGCGGCGGCGGCGGCGGCGGAGGCGGAGGCGGAGGCG 1584
Db 1381 GCGAGGAAGCGGGAAGCGGCGGCGGCGGCGGCGGCGGAGGCGGAGGCGGAGGCG 1584
QY 1585 CCGCGCGCCCAAGAGCCAGCAGTGGAGGCGCTGGGACCGCGGAGGCGGAGGCGT 1644
Db 1400 CCGCGCGCCCAAGAGCCAGCAGTGGAGGCGCTGGGACCGCGGAGGCGGAGGCGT 1459
QY 1645 CAGGACCCCGCGGAGGCGCTCTGGGAGTGTCTGGGACAGCCGCGGAGGCGGAGGCGT 1704
Db 1460 CAGGACCCCGCGGAGGCGCTCTGGGAGTGTCTGGGACAGCCGCGGAGGCGGAGGCGT 1519
QY 1705 GCAGCAGGCTCAGTGGCAGGACCGGAGGATCACCACCGCGGAGGCGGAGGCGTCA 1764
Db 1520 GCAGCAGGCTCAGTGGCAGGACCGGAGGATCACCACCGCGGAGGCGGAGGCGTCA 1579
QY 1765 CTTTCCAGAGTGAGAGATGAAGGCGATGAAGGAGTGTCTGTGGCGCCACCAAGATCAACT 1824
Db 1580 CTTTCCAGAGTGAGAGATGAAGGAGTGTCTGTGGCGCCACCAAGATCAACT 1639
QY 1825 CGAGCGCCATCAAGTGCACCTCAGCGCACAGTGCAGTGCAGATGAAGAAGCAGAAAG 1884
Db 1640 CGAGCGCCATCAAGTGCACCTCAGCGCACAGTGCAGTGCAGATGAAGAAGCAGAAAG 1699
QY 1885 TGTCCACCCCTAGTACATACACTCTGTCTTCTCAAGGCGGAGGCGGAGGCGCTCTGAA 1944
Db 1700 TGTCCACCCCTAGTACATACACTCTGTCTTCTCAAGGCGGAGGCGGAGGCGCTCTGAA 1759
QY 1945 CTACTGGGAGCTTCGAGCGCTGTGGGAGCCAGGCTCCG-CCTTAGTCCCGCAACTCT 2003
Db 1760 CTACTGGGAGCTTCGAGCGCTGTGGGAGCCAGGCTCCGCGCCCTAGTCCCGCAACTCT 1819
QY 2004 GAGCCCATGTTCTGCCCGCCAGCCCAAGGAGGAGGCGCTCAGCTTACCCAAACCTTAG 2063
Db 1820 GAGCCCATGTTCTGCCCGCCAGCCCAAGGAGGAGGAGGCGCTCAGCTTACCCAAACCTTAG 1879
QY 2064 TTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 2123
Db 1880 TTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 1939
QY 2124 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATCCCGCGGCGCAATGGA 2183
Db 1940 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATCCCGCGGCGCAATGGA 1999
```

```
QY 2184 AACCTCACGCTGGGCTCTAATTACCAGTCTTTAAAGGCCAGCCCTAGAAACCCAGCT 2243
Db 2000 AACCTCACGCTGGGCTCTAATTACCAGTCTTTAAAGGCCAGCCCTAGAAACCCAGCT 2059
QY 2244 CTTCTCTGG-AACCGCTACCTAGAGCCAGACCAAGTTACTCAGGGCTCTCCCGAGCTT 2302
Db 2060 CTTCTCTGGAAACCGCTACCTAGAGCCAGACCAAGTTACTCAGGGCTCTCCCGAGCTT 2119
QY 2303 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGGACAGGTTCCCACTCCAGCCGGG 2361
Db 2120 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGGAAANGTCCCACTTCAGCCGGG 2178

RESULT 4
HSURF HSURF 1837 bp mRNA PRI 03-JUL-1998
LOCUS H.sapiens mRNA 3'-region (unknown function).
DEFINITION Y12338
ACCESSION Y12338
VERSION Y12338.1 GI:2217971
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1837)
AUTHORS Kedra,D.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-1997) D. Kedra, Karolinska Hospital, Department
OF Molecular Medicine, Building L-6, S-171 76 Stockholm, SWEDEN
REFERENCE 2 (bases 1 to 1837)
AUTHORS Kedra,D., Seroussi,E., Fransson,I., Trifunovic,J., Clark,M.,
Lagercrantz,J., Blennow,E., Mehlin,H. and Dumaniski,J.
TITLE The germinal center kinase gene and a novel CDC25-like gene are
located in the vicinity of the PGM gene on 11q13
JOURNAL Hum. Genet. 100 (5-6), 611-619 (1997)
MEDLINE 98001089
FEATURES
source
1..1837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11q13"
<!.1027
/function="unknown"
/feature="unnamed protein product"
/codon_start=2
/protein_id="CAA73007.1"
/db_xref="GI:2217972"
/translation="LGHLERYPMALGNLADLEELPTGPRPDLFLYHKGTASAKTYT
RDEHIYPYMLAGYHCRNRNVRALQAWDATVTDYNYCREDEEIKYKFEVANDY
IPNLKEAASLLEAGEERPEQSGQSGSALQDPECFAHLRLFYDGTCKWEGSPT
PVLHVWATFLVQSLGRPEGOVRKRVISREAEAAEAEPEWGEAREGRRRPARRS
KPEPPPPKPKALDKLCTGOGAVSGPRKPGTAVGTARGPEGGTAQVAPAPASPP
PEGPVLTFQSEKMKMKELLVATKINSSAIKLQLTAQSQVQMKKQKVPSPDYLTLFL
RRQRKGL"
BASE COUNT 412 a 599 c 500 g 326 t
ORIGIN

Query Match 64.2%; Score 1781; DB 97; Length 1837;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 0; Indels 5; Gaps 5;

QY 917 CTTGGGACATCTGGAAGGTACCCCATGCCCTTAGGAACTGGCAGATCTAGAGAGCT 976
Db 1 CTTGGGACATCTGGAAGGTACCCCATGCCCTTAGGAACTGGCAGATCTAGAGAGCT 60
QY 977 GGAGCCCAACCCCTGGCGGCGGAGAGCCACTCACCTCTACCACAGGCGATTCGCTCAGC 1036
Db 61 GGAGCCCAACCCCTGGCGGCGGAGAGAGCCACTCACCTCTACCACAGGCGATTCGCTCAGC 120
QY 1037 CAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 1096
Db 121 CAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 180
```

QY 1097 TCGCAACCGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGCGCCACTGTCTATCCA 1156
DB 181 TCGCAACCGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGCGCCACTGTCTATCCA 240
QY 1157 GGACTACAACCTACTGCGCGGAAGACGAGGAGATCTACAAGAGTTCTTTTGAAGTAGCCAA 1216
DB 241 GGACTACAACCTACTGCGCGGAAGACGAGGAGATCTACAAGAGTTCTTTTGAAGTAGCCAA 300
QY 1217 TGATGTCATCCCAACCTGCTGAAGGAGGAGCAGCCAGCTTCTGAGGCGCGGAGGAGCG 1276
DB 301 TGATGTCATCCCAACCTGCTGAAGGAGGAGCAGCCAGCTTCTGAGGCGCGGAGGAGCG 360
QY 1277 GCGGGGAGCAAAAGCCAGGCGCCAGGAGCCAAAGTTTCCGCCCTCCAGGACCCCTGAGTG 1336
DB 361 GCGGGGAGCAAAAGCCAGGCGCCAGGAGCCAAAGTTTCCGCCCTCCAGGACCCCTGAGTG 420
QY 1337 CTTGCGCCACTGCTGGCATTTCTACGACGCGATCTGCAATGGAGGAGGCGAGTCCAC 1396
DB 421 CTTGCGCCACTGCTGGCATTTCTACGACGCGATCTGCAATGGAGGAGGCGAGTCCAC 480
QY 1397 GCCTGTCTGCAGTGGGCTGGGCGCCACTTTCTGTGTCAGTCCCTAGGCCGTTTGGAGG 1456
DB 481 GCCTGTCTGCAGTGGGCTGGGCGCCACTTTCTGTGTCAGTCCCTAGGCCGTTTGGAGG 540
QY 1457 ACAGGTGCGGCAGAGGTGCGCATAGTGTAGCCGAGAGCGCGGCGGCGGAGGCGCGAGGA 1516
DB 541 ACAGGTGCGGCAGAGGTGCGCATAGTGTAGCCGAGAGCGCGGCGGCGGAGGCGCGAGGA 600
QY 1517 GCGGTGGGCGAGGAGCCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1576
DB 601 GCGGTGGGCGAGGAGCCGCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 1577 AGAGGACCCCGCGCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1636
DB 661 AGAGGACCCCGCGCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 1637 TGCAGTGTACGAGGACCCCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1696
DB 721 TGCAGTGTACGAGGACCCCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 1697 TGAAGTGGCAGCAGCGCTCAGTGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1756
DB 781 TGAAGTGGCAGCAGCGCTCAGTGTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 1757 AGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGCGATGCTGTGGCGACCAA 1816
DB 841 AGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGCGATGCTGTGGCGACCAA 900
QY 1817 GATCAACTGAGCGGCCATCAAGCTGCAACTCAGGCGACAGTTCGCAAGTGCAGATGAAGAA 1876
DB 901 GATCAACTGAGCGGCCATCAAGCTGCAACTCAGGCGACAGTTCGCAAGTGCAGATGAAGAA 960
QY 1877 GCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTTCTCAAGCGGCGGCGGCGGCGGCGG 1936
DB 961 GCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTTCTCAAGCGGCGGCGGCGGCGGCGG 1020
QY 1937 CCTCTGAACCTACTGGGACTTGGGACCGCTTGTGGGACCGAGGCTCCG-CCTTAGTCCC 1995
DB 1021 CCTCTGAACCTACTGGGACTTGGGACCGCTTGTGGGACCGAGGCTCCGCGCTTAGTCCC 1080
QY 1996 CCAACTCTGAGCCCATTTCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2055
DB 1081 CCAACTCTGAGCCCATTTCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 2056 ACCCTAGTGTCCCGGCTCCGAGTACCTGTGTATCAACCCAGATTTCTCCAGGCTCAG 2115
DB 1141 ACCCTAGTGTCCCGGCTCCGAGTACCTGTGTATCAACCCAGATTTCTCCAGGCTCAG 1200
QY 2116 AACCCAGGCTCTGCGCCCGGCTGAGTATAGGCTCTTCTTCCCGGAGATCCAGCGCGG 2175
DB 1201 AACCCAGGCTCTGCGCCCGGCTGAGTATAGGCTCTTCTTCCCGGAGATCCAGCGCGG 1260

QY 2176 CCAATGAAACCTCAGCTGGGTCTCTAATTTACAGTCTTTTAAAGGCGCCAGCCCTAGAAA 2235
DB 1261 CCAATGAAACCTCAGCTGGGTCTCTAATTTACAGTCTTTTAAAGGCGCCAGCCCTAGAAA 1320
QY 2236 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTTACTCAGGGCTCCCTC 2295
DB 1321 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTTACTCAGGGCTCCCTC 1380
QY 2296 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTTACTCAGGGCTCCCTC 2355
DB 1381 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTTACTCAGGGCTCCCTC 1439
QY 2356 CCGGGGAGCTTAGGACCACTCAGCCCTCAGGAGTATATTTCCGCACTTCCAGAAATTCAT 2415
DB 1440 CCC-GGAGGCTTAGGACCACTCAGCCCTCAGGAGTATATTTCCGCACTTCCAGAAATTCAT 1498
QY 2416 ATCTTGGAAATCAAGCTCCTGCTGATCCGCAATAACTTCAGTCTCTTCCAGAAATTCGAA 2475
DB 1499 ATCTTGGAAATCAAGCTCCTGCTGATCCGCAATAACTTCAGTCTCTTCCAGAAATTCGAA 1557
QY 2476 ATCTTGGAAATCAAGCTCCTGCTGATCCGCAATAACTTCAGTCTCTTCCAGAAATTCGAA 2535
DB 1558 ATCTTGGAAATCAAGCTCCTGCTGATCCGCAATAACTTCAGTCTCTTCCAGAAATTCGAA 1617
QY 2536 TGAGCTCTCTGAGCCCGGCTCTCTCTGAGCAAACTGGCCCGGATCAGAGCAGGACCT 2595
DB 1618 TGAGCTCTCTGAGCCCGGCTCTCTCTGAGCAAACTGGCCCGGATCAGAGCAGGACCT 1677
QY 2596 CCCTTCGAGCCCTCTGGGAACCTCCAGAGGTCGAGGAGCTCCGAGCATCCCGGAGGA 2655
DB 1678 CCCTTCGAGCCCTCTGGGAACCTCCAGAGGTCGAGGAGCTCCGAGCATCCCGGAGGA 1737
QY 2656 AATCTGAGAGGGTGTAGGAGTGGTGACAAAGAGCTGATCTCTCTCTCTCTCTCTCTCTCT 2715
DB 1738 AATCTGAGAGGGTGTAGGAGTGGTGACAAAGAGCTGATCTCTCTCTCTCTCTCTCTCT 1796
QY 2716 GATTTATTTTCTAGTTCCCAAGAAAGATGAATATTTTGT 2756
DB 1797 GATTTATTTTCTAGTTCCCAAGAAAGATGAATATTTTGT 1837

RESULT 5

AB023400 3067 bp mRNA ROD 02-MAR-2000
LOCUS Rattus norvegicus MEN1 mRNA for menin, complete cds.
DEFINITION
ACCESSION AB023400
VERSION AB023400.1 GI:5360728
KEYWORDS MEN1; menin.
SOURCE Rattus norvegicus (strain:Wistar) Brain cDNA to mRNA.
ORGANISM Rattus norvegicus
REFERENCE 1 (sites)
AUTHORS Maruyama, K., Tsukada, T., Hosono, T., Ohkura, N., Kishi, M., Honda, M., Nara-Ashizawa, N., Nagasaki, K., and Yamaguchi, K.
TITLE Structure and distribution of rat menin mRNA
JOURNAL Mol. Cell. Endocrinol. 156 (1-2), 25-33 (1999),
MEDLINE 20077667
REFERENCE 2 (bases 1 to 3067)
AUTHORS Maruyama, K., Tsukada, T., Nagasaki, K., Maruyama, K., Tsukada, T., Nagasaki, K., Ohkura, N., Honda, M., Nara, N., Hosono, T., Kishi, M., Kazuki, S., Futami, H., and Yamaguchi, K.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-1999) to the DDBJ/EMBL/GenBank databases. Kouji Maruyama, National Cancer Center Research Institute, Growth Factor Division; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(E-mail: kmaruyamegan2.res.ncc.go.jp, tel:81-3-3542-2511(ex.4302), Fax:81-3-3542-8170)
FEATURES
source
1. .3067
/organism="Rattus norvegicus"
/strain="Wistar"


```

/db_xref="taxon:10116"
/tissue_type="Brain"
470..2302
/gene="MEN1"
470..2302
/gene="MEN1"
/codon_start=1
/product="menin"
/protein_id="BAA82134.1"
/db_xref="GI:5360729"
/translation="MGLRAAKTKLPLRSIDVVLPAELGPREPDLVLSLVGFV
EHLAVNVIPTNPELVTFQSPADPPGLTYFPVADLSIAALYARFTAIRGAVD
LSLYPREGVSRELKVSQSDYIWNLSRYSFKDRHIQSLSFSETITQKLDSSGVAF
VWGCAQLGLRDVHLALSEDHAWVFGNSGQTAETVWHGKQNEDRQOTVNAVAER
SVLYLKGSYMRCDRKMVAVMCAINPSIDILSDLSLELLQLQKLLWLLYDLHLERY
PMALGNLADLELEPTPCRPDPLLYHKGIASAKTYODEHYIPWYLAGYHCRNRV
REALQWADTATYIQDYNCREDEIYKEFFEVSANDVIPNLLKEAASLLEAGERPE
QAQGTGGADQLQPECAHLLRYDGIKWEESGPTVLHVGNATFLVLSGRFEGQ
VRQVHIVSRSEAEEAEPEWGEAREGRPRRESKPEPPPPKPKALDKPGSGQ
SAGSPPRKTSCTVSGTARGTEVSAAQAPAPASPPPEPGLVTFQSEKMKMKELLY
ATKINSSAIKQLQTAOSQVMKKQKVSPTSDYTLFLKQRKGL"
BASE COUNT      672 a  904 c  881 g  610 t
ORIGIN

Query Match      59.5%; Score 1648.6; DB 94; Length 3067;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 519; Indels 71; Gaps 11;

QY  43  GGAGGGCCCCGGCCCACTATTTCAGGCTCTGGGGGAGGGCCGCCGCCACCGC 102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  402  GGAGGGGGCTCTCGCCGAGCCGACCCCTCCCTTCCCTGCGCTTGCAGCCACCGC 461

QY  103  CGCGCCCATGGGCTGAAGCGCCGCCAGACGCTGTTCGGCTGCGCTCCATCGAGC 162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  462  CGCGCCCATGGGCTGAAGCGCCGCCAGACGCTGTTCGGCTGCGCTCCATCGAGC 521

QY  163  AGTGTGCGCGCTGTTCGTCGAGCTGGCGCGAGAGGCGGACCTTGGTGTCTCTT 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  522  AGTGTGCGCGCTGTTCGTCGAGCTGGCGCGAGAGGCGGACCTTGGTGTCTCTT 581

QY  223  CTTTGTGTCGGGCTTGTGGAGCATTTTTCGTGTCTAACCGCGTCACTCCATACCAAG 282
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  582  CTTTGTGTCGGGCTTGTGGAGCATTTTTCGTGTCTAACCGCGTCACTCCATACCAAG 641

QY  283  TTCCCGAGCTCACTTCCAGCCAGCCCGCCCGCCCGCCCGCTGCGCGCTCACTACT 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  642  TGCCGGAGCTCACTTCCAGCCAGCCCGCCCGCCCGCCCGCTGCGCGCTCACTACT 701

QY  343  TTCCCGTGCGGACCTGTCTATCATCGCGCCCTCTATCGCGCTTTCACCGCCAGATCC 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  702  TCCAGTGCGGACCTATCATCATTTGTCGCTCTATCGCGGATTCACCGCTCAGATCC 761

QY  403  GAGGGCGCGTCGACCTGTCCCTCTATCTCTCGAGAAGGGGTGTCTCCAGCCGTGAGCTGG 462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  762  GCGGTGCTGTGACCTCTCCCTCTATCTCTCGAGAAGGGGGGCTTCCAGTCGCGAATCG 821

QY  463  TGAAGAGTCTCCGATGTCATATGGAACAGCTCAGCGCTCTACTTCAAGGATCGGG 522
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  822  TAAAAGAGTCTCGGATGTCATATGGAACAGCTCAGCGCTCTACTTCAAGGATCGGG 881

QY  523  CCACATCCAGTCCCTCTTCACTCATCACAGGACCAAAATTTGACAGCTCCGGTGTGG 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  882  CCACATCCAGTCTCTCTTCACTCATCACAGGACCAAAATTTGACAGCTTCCGGGTGG 941

QY  583  CTTTGTCTGTTGGGCTGCCAGCCCTCGGTCTCCGGGATGTCCACCTCGCCCTGT 642
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  942  CTTTGTCTGTTAGGGCTGCCAGCTCTAGGCTCAGAGATGTCCATCTGGCCCTGT 1001

QY  643  CTGAGATCATGCTGGGTAGTGTGTTGGGCCCAATGGGGAGAGACAGCTGAGGTCACT 702
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1002  CGAGGATCATGCTGGGTAGTGTGTTGGGCCCAATGGGGAGAGAGCTGAGGTGACCT 1061
```

```

QY  703  GGCAGCGCAAGGGCAACGAGGACCGGCGAGAGTCAATGCCGTGTCCTGAGC 762
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1062  GGCATGGCAAGGCAACGAGGACCGGCGAGAGGCGAGAGTCAACGCTGCTGCTGAGC 1121

QY  763  GGAGCTGGCTTACCTTGAAAGGATCATATCGCTGTGACCGCAAGATGGAGGTGGCGT 822
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1122  GGAGCTGGCTTACCTTGAAAGGATCATATCGCTGTGACCGCAAGATGGAGGTGGCGT 1181

QY  823  TCATGGTGTGTGCCATCAACCCCTTCCATTGACCTGCACACCGACTCGCTGAGCTTCTGC 882
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1182  TCATGGTGTGTGCCATCAACCCCTTCCATTGACCTGCACACCGACTCGCTTGGAACTTTC 1241

QY  883  AGCTGCAGCAGAGCTGCTCTGCTGCTCTATGACCTGGGACATCTGGAAGGTATCCCCA 942
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1242  AACTACAGCAGAGCTGCTCTGCTGCTGTATGACCTAGGACATCTGGAAGGTATCCCCA 1301

QY  943  TGGCTTTAGGAACTTGGGAGATCTAGAGAGCTGGAGGCCACCCCTGGCGGCCAGACC 1002
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1302  TGGCGCTAGGAACTTGGGAGACTTGGAGAGCTGGAGGCCACCCCTGGCTGGGCCAGACC 1361

QY  1003  CACTCACCTCTACCAACAGGGCATTTGCCCTCAGCCAGAGACTACTATCGGATGAACACA 1062
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1362  CGCTCACCTCTATCAAGAGGATTTGCCCTCAGCTAAGACTTACTACAGGATGAACACA 1421

QY  1063  TCTACCCCTTACATGTACTGCTGGCTACCACTGTTCGCAACCGCAATGTGCGGGAAGCCC 1122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1422  TCTACCCCTTACATGTACTGCTGGCTACCACTGTTCGCAACCGCAATGTGCTGGAAGCCC 1481

QY  1123  TCGAGGCTTGGGGGAGACGGCCACTGTCTATCCAGGACTACAACTACTTGCCTGGGAGAGC 1182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1482  TCGAGGCTTGGGGGAGACGGCCACTGTCTATCCAGGACTACAACTACTTGCCTGGGAGAGC 1541

QY  1183  AGSAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATCTATCCCAACCTGCTGAAG 1242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1542  AGSAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATCTATCCCAACCTGCTGAAG 1601

QY  1243  AGGCAAGCTTGTTCGAGGGGCGAGGAGCGCGGGGGAGCAAGAGCCAGGCAAGCC 1302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1602  AGGCAAGCTTGTTCGAGGGGCGAGGAGCGCGGGGGAGCAAGAGCCAGGCAAGCC 1661

QY  1303  AGAGCCAAAGTTTCGGCCCTCCAGGACCTTGTGCTTTCGGCCACCTGCTGCGATTCATAG 1362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1662  AGAGCCAAAGTTTCGGCCCTCCAGGACCTTGTGCTTTCGGCCACCTGCTGCGATTCATAG 1721

QY  1363  ACGGATCTGCAAAATGSGAGGAGGAGTCCCAACGCTGTGCTGCAAGTGGGCTGGGCA 1422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1722  ACGGATCTGCAAAATGSGAGGAGGAGTCCCAACGCTGTGCTGCAAGTGGGCTGGGCA 1781

QY  1423  CCTTTCTTGTGAGTCCCTTAGGCGCTTTTGAAGGACAGGTTCGGGAGAGAGTGGCATAG 1482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1782  CCTTTCTTGTGAGTCCCTTAGGCGCTTTTGAAGGACAGGTTCGGGAGAGAGTGGCATAG 1841

QY  1483  TGAGCCGAGAGCCGAGGCGCGAGGCGAGGCGGTGGGGGAGAGAGCCCGGGAAG 1542
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1842  TTAGCCGGAAGCCGAGGCGAGAGGCGGAAGAACCATATGGGGGAGAGAGCCCGGAAG 1901

QY  1543  GCGGGGCGGGGCGCCACCGCGGGAGTCCAAGCCAGAGAGGCCCCCGCCGCCCAAGAGC 1602
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1902  GTCGTGCGGCGGCTCCCGCCGAGAGTCCAAGCTTGAAGCTTGAAGGACCCACCCCAAGAGC 1961

QY  1603  CAGCACTGACAAAGGCGCTGGGCGCCGCGGAGGTGTGAGGACCCCGCCCGGAAGC 1662
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  1962  CAGCACTGACAAAGGCGCTGGGCGCCGCGGAGGTGTGAGGACCCCGCCCGGAAGC 2021

QY  1663  CTCTTGGGACTGTCTGCTGGCACAGCCCGAGGCGCTTGAAGGTGGCAGCAGGCTCAGGTGC 1722
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  2022  CGTCAAGGACTGTCTCAGGTACTTGCCTGGCACAGAGTCAAGGAGTGTGCTCAGGCTC 2081

QY  1723  CAGCACCCCGCAGCATCCACCCCGGAGGCTCAGTGTCTACTTCCACAGGTGAGAGA 1782
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  2082  CAGCACTGACAGTCCACCCCGAGGCGGCTGCTACTTCTCAGGTGAGAGGAGT 2141

QY  1783  TGAAGGGCATGAAGGAGCTGCTGTTGGGCCCAAGATCAACTCGAGCGCCATCAAGCTGC 1842
```


Db	130	GCCAAATGATGTCATCCCAACCTGCTGAAAGAGGAGCAGCCAGTTCGTGGAGCGGGCGAG	189
Qy	1272	GAGCGCGCGGGGAGCAAAAGCCAGAGGSCACCCAGAGCCAAAGTTTCGGCTCCTCCAGAGCCCT	1331
Db	190	GAGCGCGCGGGGAGCAAAAGCCAGAGGSCACCCAGAGCCAAAGTTTCGGCTCCTCCAGAGCCCT	249
Qy	1332	GAGTGTTCGCCACCTGCTGCGATTCTACGACGCGATCTGCAAAATGGGAGGAGGGCAGT	1391
Db	250	GAGTGTTCGCCACCTGCTGCGATTCTACGACGCGATCTGCAAAATGGGAGGAGGGCAGT	309
Qy	1392	CCACAGCCTGCTGCTGACGCTGGGCTGGGCCACTTTCTTGTCAGTCCCTAGGCGCGTTTT	1451
Db	310	CCCAGCGCTGCTGCAAGTGGGCTGGGCCACTTTCTTGTCAGTCCCTAGGCGCGTTTT	369
Qy	1452	GAGGACAGGTGCGGCAGAAAGTGGCGCATAGTGAAGCCAGAGGCGCGAGGGCGCGAGGCC	1511
Db	370	GAGGACAGGTGCGGCAGAAAGTGGCGCATAGTGAAGCCAGAGGCGCGAGGGCGCGAGGCC	429
Qy	1512	GAGGAGCGTGGGGCGAGGAAGCCCGGGAAGCCCGGCGGGGCGCCACAGCGCGGAGTCC	1571
Db	430	GAGGAGCGTGGGGCGAGGAAGCCCGGGAAGCCCGGCGGGGCGCCACAGCGCGGAGTCC	489
Qy	1572	AAGCCAGAGAGAGCCCGCGGCCACAGAGCCAGCAGTGGACAAAGGGCCTGGGCAACCGGC	1631
Db	490	AAGCCAGAGAGAGCCCGCGGCCACAGAGCCAGCAGTGGACAAAGGGCCTGGGCAACCGGC	549
Qy	1632	CAGGGTCAAGTGTACAGAGCCCGCGGAAGCTCCTGGGACTGTGCTGGCGACAGCCCGA	1691
Db	550	CAGGGTCAAGTGTACAGAGCCCGCGGAAGCTCCTGGGACTGTGCTGGCGACAGCCCGA	609
Qy	1692	GGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGAGCATCACACCGCGGAG	1751
Db	610	GGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGAGCATCACACCGCGGAG	669
Qy	1752	GGTCCAGTCTCACTTTCAGAGTGAAGAATGAAGGCATGAAGGAGCTGCTGGTGGCC	1811
Db	670	GGTCCAGTCTCACTTTCAGAGTGAAGAATGAAGGCATGAAGGAGCTGCTGGTGGCC	729
Qy	1812	ACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACAGTCGCAAGTGCAGATG	1871
Db	730	ACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACAGTCGCAAGTGCAGATG	789
Qy	1872	AAGAAGCAGAAAGTGTCCACCCCTAGTCACTACACTGTCTTTCCCTCAAGCGGACGCG	1931
Db	790	AAGAAGCAGAAAGTGTCCACCCCTAGTCACTACACTGTCTTTCCCTCAAGCGGACGCG	849
Qy	1932	AAAGGCTCTGAACACTCTGGGACCTTCGGACCGCTTGTGGGAGCCAGGCTCCG - CTTTA	1990
Db	850	AAAGGCTCTGAACACTCTGGGACCTTCGGACCGCTTGTGGGAGCCAGGCTCCGCGCTTA	909
Qy	1991	GTCCCCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGAGGAGCAGGCTCACCTCTA	2050
Db	910	GTCCCCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGAGGAGCAGGCTCACCTCTA	969
Qy	2051	CCCAAGCCCTAGTTCCTCCGCTCCGAGTACGTCGTATCAAAACCCACGATTTTCTCCAG	2110
Db	970	CCCAAGCCCTAGTTCCTCCGCTCCGAGTACGTCGTATCAAAACCCACGATTTTCTCCAG	1029
Qy	2111	CTCAGAACCCAGGGCTCTGCCCGAGTGGTTAGAATATAGTGTCTTCTCCAGAAATCCCA	2170
Db	1030	CTCAGAACCCAGGGCTCTGCCCGAGTGGTTAGAATATAGTGTCTTCTCCAGAAATCCCA	1089
Qy	2171	GCSCGCCAATGAAACCTCACGCTGGTCTTAATTAACAGTCTTTAAAGGCCAGGCCCT	2230
Db	1090	GCSCGCCAATGAAACCTCACGCTGGTCTTAATTAACAGTCTTTAAAGGCCAGGCCCT	1149
Qy	2231	AGAAACCCAAAGCTCTCTCCGGAACCGCTTCAGTACAGCCAGACCAACGTTACTCAGGGC	2290
Db	1150	AGAAACCCAAAGCTCTCTCCGGAACCGCTTCAGTACAGCCAGACCAACGTTACTCAGGGC	1209
Qy	2291	TCCTCCAGCTGTAGGAGCTGAGGTTTACCTTTAACCCAAAGGGAGCAGAGTCCCAAC	2350

[illegible]

/note="similar to Homo sapiens MENIN"

/codon_start=1

/product="menin"

/db_xref="GI:5006430"

/translation="MGKAAQKTLFPLRSIDVVRLFAELGREBPDVLVLLSLVLFV
EHLAVRNVIPTNPVLETPSPAPDPGLTFPPVADLSIIAALYARETAQIRGAVD
LSLYREGVSSRELKYSVDVWNSLSRSYKFDRAHIOQLSFITGKLDSSGVAFA
VGCACALGLRDVHLALSDEHAWFVPCNGEOTAEVTHGKNGEDRRGOTVNAQVAFR
SWLYLAKSYMRCRDMKMEAFMCAINPSIDLTHTDSLELLOQLKLLVLDLGLHLYR
PHALGNLADLELEPTPGRPDLTLVHKGIASAKTYQDEHYLPYMWLAGYHCRNRV
REALQAWDTATVQIDNFCREDEEYKEFFEVANDVIPNLLKEASLLEETGEERGE
QAQGTGGQSAQDPFCFAHLRLFDGICKWEEGSPVPLHVGNATFLVQSLGFEQG
VRQKHIVREAAAEAEPEWGEARERRRGPRESKPEEPKPPALDRGPGSGQ
SAGSGPPRTSGTVPGTTRGGQEVGNAAPAPAAAPPPPEGPVLTTFQSEKMGMKELL
VATKINSSAIKILQTAQSOVMKKOKVSPSDYTLSELKRQRKL"

BASE COUNT 656 a 889 c 859 g 580 t

ORIGIN

Query Match 58.3%; Score 1616.4; DB 94; Length 2984;

Best Local Similarity 79.1%; Pred. No. 4.8e-294;

Matches 2130; Conservative 0; Mismatches 446; Indels 116; Gaps 13;

QY	88	GCCGCCGCCACCGCCGCCCGCCGATGGGCTGAAGCGCCGCCAGAGACGCTGTTCCTCCG	147
DB	399	GCCGCCAGCCACCGCCGCCCGCCGATGGGCTGAAGCGCCGCCAGAGACGCTGTTCCTC	458
QY	148	TGCCTCCATCGAGAGCTGGTGGCTGTTTGTGCGGAGCTGGCGCGAGAGAGCGG	207
DB	459	TGCCTCTATCGAGAGCTGGTGGCTGTTTGTGCGGAGCTGGCGCGAGAGAGCGG	518
QY	208	ACCTGGTCTCTTTCCTTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCG	267
DB	519	ACCTGGTCTCTTTCCTTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCG	578
QY	268	TCATCCCTACCAAGTTCCGAGCTACCTTCAGCCCGCCCGCCCGCCCGCCGCTG	327
DB	579	TCATCCCGCCAGCTGCGGAGCTACCTTCAGCCCGCCCGCCCGCCCGCCGCTG	638
QY	328	CGGCGCTACCTACTTTCCTGGTGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCT	387
DB	639	GTGGCTTACCTACTTTCCTGGTGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCT	698
QY	388	TCACCGCCCGAGATCCGAGCGCCGCTGACCTGTCCCTCTATCTTCGAGAGGGGTCTCT	447
DB	699	TCACCGCTCAGATCCGCGCGCTGTGACCTCTCCCTCTATCTTCGAGAGGGGTCTT	758
QY	448	CCAGCGTGAGCTGGTGAAGAGTCTCCGATGTCTATGAGACAGCTCAGCGGCTCTCT	507
DB	759	CTAGTCGCGAATCGTAAAGAGCTCTCGATGTCTATGAGACAGCTCAGCGGCTCTCT	818
QY	508	ACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAGCTTCATCACAGGACCAATTTG	567
DB	819	ACTTCAAGGACCGGCGCCACATCCAGTCCCTCTTCAGCTTCATCACAGGACCAATTTG	878
QY	568	ACAGTCCGCTGTGGCTTTCTGTGTGGGCTGCCAGCGCCCTGGGCTCTCCGGATG	627
DB	879	ACAGTCCGCGCTGGCTTTGCACTGGTGGGCTGCCAGGCTCTGGTCTCAGAGATG	938
QY	628	TCCACTTCGGCTCTGTAGAGATCATGCTGGGTAGTGTGTGGGCCCAATGGGGAGAGA	687
DB	939	TCCATCTGGCGCTGTCTGAAGATCATGCTGGGTGGTGTGTGGGCCCAAGGGGAGAGA	998
QY	688	CAGCTGAGGTCACTGCGACGGCAAGGCAACAGAGACCGCGGCGCCAGACAGTCAATG	747
DB	999	CAGCTGAGGTGACGTGGCAACGGCAAGGCAACAGAGACCGCGGCGCCAGACAGTCAATG	1058
QY	748	CCGCTGTGGCTGAGCGGAGTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA	807
DB	1059	CCGCTGTGGCTGAGCGGAGTGGCTGTACCTGAAAGGAGTGTGTATCGCTGGACCGTA	1118
QY	808	AGATGGAGGTGGCGCTTCATGGTGTGGCCATCAACCCCTTCATTTGACCTGCACACCGACT	867

DB	1119	AGATGGAGGTGGCGTTTCATGGTGTGCCATCAACCCCTTCATCGATCTTTCACACTGACT	1178
QY	868	CGCTGGAGCTTTCGACGCTCAGCAGAAAGCTGCTGCTGGCTGCTTATGACCTGGGACATC	927
DB	1179	CTTTGGAACTGTTTGCACACTACAGCAGAAAGCTGCTGCTGGCTGCTGTATGACCTGGGACATC	1238
QY	928	TGAAAGGTACCCCATGGCCCTTAGGAAACCTGGCAGACTTAGAGGAGCTGGAGCCACCC	987
DB	1239	TGAAAGATACCCCATGGCCCTTAGGAACTTGGCAGACTTGGAGGAGCTGGAGGCTACCC	1298
QY	988	CTGGCGGCGAGACCCACTCACCTCTACCAAGGCGCTTGCCTCAGCCAAAGACCTACT	1047
DB	1299	CGGCGCGGCGAGACCCACTCACCTTATCACAAGGAAATTTGCTCAGCTAAGACCTACT	1358
QY	1048	ATCGGATGAACACATCTACCCCTTACATGTACCTGGCTGCTACCACTGTGCAACCGCA	1107
DB	1359	ACCAGGATGAACACATCTACCCCTTACATGTACCTGGCTGCTACCACTGTGCAACCGCA	1418
QY	1108	ATGTGCGGGAAGCCCTGAGGCTTGGCGGACAGCGGCACTGTCTATCCAGGACTTACAAC	1167
DB	1419	ATGTGCGGGAAGCCCTGAGGCTTGGCGGACAGCGGCACTGTCTATCCAGGACTTACAAC	1478
QY	1168	ACTGCGGGAAGAGGAGATCTACAGGAGTTCCTTTGAAGTAGCCCAATGATGTCTATCC	1227
DB	1479	ACTGCGGGAAGGAGATCTACAGGAGTTCCTTTGAAGTAGCCCAATGATGTCTATCC	1538
QY	1228	CAAACCTGCTGAAGGAGCGACGCTTGTGAGGCGGCGGAGGCGCGCGGGGAGC	1287
DB	1539	CAAACCTGCTGAAGGAGCGGCGGAGGCTGTCTAGAACAGGCGGAGGAGCTGGGGAGC	1598
QY	1288	AAAGCCAGGCGACCCAGAGCAAGGTTCCGCTTCCAGGACCTGTAGTGTCTCCGCCACC	1347
DB	1599	AAGCCAGGCGACCCAGGCGCAAGTTCCTCTCTCCAGGACCCAGAGATGCTTTGCCACC	1658
QY	1348	TGCTGCGATTTCTAGGAGCGGATCTGCAAAATGGGAGGAGGCGAGTCCACACGCTGTCTGC	1407
DB	1659	TGCTGCGATTTTATGATGGCATCTGCAAAATGGGAGGAGGCTAGCCACACGCTGTCTGC	1718
QY	1408	ACGTGGCTGGCGCACCTTTCTGTGAGTCCCTTAGGCGCTTTTGAGGGACAGAGTGGCGG	1467
DB	1719	ACGTAGGCTGGCGCACCTTTCTGTGAGTCCCTTAGGCGCTTCGAGGGACAGAGTGGCGG	1778
QY	1468	AGAAGTGGCGATAGTGAAGCGGAGGCGGCGGCGGCGGAGGAGGCGGCTGGGGGCG	1527
DB	1779	AGAAGTGGCGATAGTGAAGCGGAGGCGGCGGCGGAGGCTGAAGAACCAGTGGGGG	1838
QY	1528	AGNAGCGCGGAGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGAGGAGCGCC	1587
DB	1839	ATGAAGCGCGAGAGGCGGCTGGCGGCTGCTCCCGAGAGAGTCCAGGCTTGAGGAGCGAC	1898
QY	1588	CGCGCGCAAGAGCGGAGCTTGACAAAGGCGCTTGGCGACCGCGGCGGCTGCACTGTCTAG	1647
DB	1899	CACCAAGAGAGGCTGCAATGGACAAAGGCGGCGGCGCTCAGGACAAAGTGCAGGCTCGG	1958
QY	1648	GACCGCGCGGAGCGCTCTCTGGAGTCTGCTGGCACAGCGCC---GAGCGCTGAAGGTG	1704
DB	1959	GACCACTTGGGAGGAGCTCAGGAGTCTCCAGGACTGTCCAGGCTTACCGCGCTGGCGAGGAGT	2018
QY	1705	GCAGCGGCTCAGTGGCGGAGCGGCGGCGGCGGAGTCCCAAGCCAGAGGAGGCTGCTCA	1764
DB	2019	GCAATGCTGTCTAGGCTTCCAGGAGCTGCGGAGGAGTCCAGGAGGAGGCGGCTGCTCA	2078
QY	1765	CTTTCCAGAGTGAAGATGAAGGAGTGAAGGAGTGTGGTGGCGGCGGCGGAGTCAACT	1824
DB	2079	CTTTCCAGAGTGAAGATGAAGGAGTGAAGGAGTGTGGTGGCGGCGGCGGAGTCAACT	2138
QY	1825	CGAGCGGCTCAAGCTCAACTCACGCGAGTGTGGAGTGTGCAAGTGAAGAGCAGAAAG	1884
DB	2139	CGAGCGGCTCAAGCTCAACTCACGCGAGTGTGGAGTGTGCAAGTGAAGAGCAGAAAG	2198
QY	1885	TGTCCACCGCTTGTGAGTACACTGTCTTCTTCTCAAGGCGGCGGCGGAGGCTGTGAA	1944

Db 2199 TGTCCACACCCAGCGACTACACACTCTCTTTCTTAAGCGGAGCGACGAGGCTCTGA - 2257
QY 1945 CTACTGGGACTTGGACCGCTTGTGGGACCGCAGGCTCCGCTTAGTCCGCCAACCTGTG 2004
Db 2258 -----ACTCTGATTTCTGTGGACGCCCGCCGCTTACTCTCCAGCAGGCTT 2308
QY 2005 AGCCCATGTTTCCGCCCGCCAGCCAAAGGGAGAGGCTTACCTCTACCCAAACCCCTAGT 2064
Db 2309 CACCTCATCTCTTAATCAAGCCTA-GACTCCGCCCTCAGTCTAGGGAGCCAGAGCT 2367
QY 2065 TCCCGGTCGGAGTACAGTCTGTATCAACACCCAGATTTTCTCCAGCTCAGAACCCAGG 2124
Db 2368 GTCCGCACTGACTA-----GGAGCCTTCAATTTAGCCCTCTTCAGGCCCCAGAG 2417
QY 2125 CTCTGCCCGAGTCGTAGATATAGTCTCTTCTCCAGAAATCCAGCGCGCAATGGAA 2184
Db 2418 CCTGGCTCAGTCTTAGAATGTAGTGTATCCCG-----ACAGCCGCGCAATTTGA 2471
QY 2185 ACTCAGCGTGGTCTTAATACAGTCTTTAAAGGCCCGCCCTTAGAAACCCCAAGCTC 2244
Db 2472 ACCTCATTTAGTCCATCCCGAAATTCAGTGGACCATGAC--AGGAGCCCAAGTCA 2529
QY 2245 CTCTCGGAACCGCTCACCTAGAGCAGACCAAGTTTACTCAGGGCTCTCCAGGTTT 2304
Db 2530 CGCCTGAACAGTAGGCACTTCTCCCACTTGGGGCTAGATTGACCTTAGCCCC----- 2582
QY 2305 AGGAGCTAGGTTTCAACCTTAACCCAAAGGAGCAGCAGCTCCACCTCCAGCCCGGGAG 2364
Db 2583 -----CAGAGCACAGTTGTACTT-----GGAAA 2606
QY 2365 CTTAGGACACTCAGCCCTTAGGAGTATATTTCCGCACTT-----CAGAATTCATATCT 2420
Db 2607 CCAAGACCAATTCAGTCCCTTAGGGTAGATTTCGCACTTTATATAGAGTTTCCATCTTT 2666
QY 2421 CGAATCCAGCTCCCTGCCCCAATACTTCAGTCTCTCTCCAGATTTTGAATCT 2480
Db 2667 GGAATCCAAATCCCTTACCCCAAGAACTTACATATTCTGCCCCGAATTTGGAATC-- 2724
QY 2481 AGTTTCTCTCTCTGATATCCGAGTCTGGGACACAAAACTCCGCCCGCCAGCTATGAGC 2540
Db 2725 ---TTCTCTGCTCTCCCTAGCCCTGAGTACAGGACATGAGCTCCGACTTATCTGTGA 2781
QY 2541 ATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCCGGATCAGAGCAGGACCTCCCTT 2600
Db 2782 ATACTGAGCTCGCCCTCTCTGACCAAACTGGCGCAAGAACAGAGCAGGACCT--CTT 2839
QY 2601 CGAGCCTCTGGAACTCCAGAGGTCAGCCCATCTCGGAGCATCCCGGAGGAATCT 2660
Db 2840 CGAGCCTCTGGAACTCCCTGAGGTC-----CCAGGAGGAATCT 2881
QY 2661 GCAGAGGGTTAGGAGTGGGTGACAAGAGCCTGATCTCTCTCTTTTGTACATAGATT 2720
Db 2882 GCACGGGTGGGGAG-----GGCCAGCCTCTCTCTTTTGTACATAGATT 2929
QY 2721 ATTTTTCAGTCCAGAAAGATGAATACATTTTGTAAAAAATAAAAAA 2772
Db 2930 ATTTTTCAGTCCAGAAAGATGAATGATTTTGTAAAAAATAAAAAA 2981

RESULT 8
AF016398
LOCUS AF016398 2623 bp mRNA ROD 30-NOV-1998
DEFINITION Mus musculus menin (Men1) mRNA, complete cds.
ACCESSION AF016398
VERSION AF016398.1 GI:3929517
KEYWORDS menin.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2623)
AUTHORS Stewart,C., Parente,F., Piehl,F., Farnebo,F., Quincey,D.,
Sillins,G., Bergman,L., Carle,G.F., Lemmens,I., Grimmond,S.,

Xian,C.Z., Khodel,S., Teh,B.T., Lagercrantz,J., Siggers,P.,
Calender,A., Van de Vem,V., Kas,K., Weber,G., Hayward,N.,
Gaudray,P. and Larsson,C.
Characterization of the mouse Men1 gene and its expression during
development
Oncogene 17 (19), 2485-2493 (1998)
REFERENCE 2 (bases 1 to 2623)
AUTHORS Stewart,C.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1997) Cancer Unit, Queensland Institute of
Medical Research, 300 Herston Road, Herston, Queensland 4053,
Australia
FEATURES
source 1..2623
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..2623
/gene="Men1"
CDS 62..1897
/gene="Men1"
/codon_start=1
/product="menin"
/protein_id="AAC79938.1"
/db_xref="GI:3929518"
/translation="MGLKAAOKTLPLRSIDVRLFAAELGREPDLVLISLVLFV
EHLAVNRVPTNPVLPETPSPADPPGLTVFPVADLSIIAALYARETAOIRGAVD
LSLYPREGVSSRELVKVSDVINWSLSRSFKDRAHTQSLSFTTGTGKLDSSGVAFA
VVGACALGRDVLHLSLSEDAHWVFGNPGNEQTAETVNHGKNDRRGOTVWAGVAER
SWLYLKGYMRCDRMEVAFMVCAINPSIDLTLSLELQLOKLLWLLYDLHLRLRY
PMALGNLADLEPTPGRPDPLTYLHKGIASAKTYQDEHLYPYMYLAGYHCRNRV
REALQAWADTATVIODYNVCREDEITYKEFFEVDVIVPNLLKEASILELTCERTGE
QAQGTGGQSALQDEPCAHLLRFYDGLCKWEEGSPTPVLHVGMATFLVQSLRGEGQ
VRQKVHIVSREAAEAEEPPWGDDEARRRRGPRRESKPEEPPEPPPKKPAIDKPGSGQ
SAGSGPPPKTSGTPTGTGGQEVNAAQAPAPAPAPPEPGLVTFQSEKMGKMKELL
VATKINSSAIKLQLTAQSQVMKKQKVSPSDYTLISFLKRRKRL"
2608
polya_site /gene="Men1"
BASE COUNT 594 a 794 c 721 g 514 t
ORIGIN
Query Match 58.3%; Score 1616; DB 94; Length 2623;
Best Local Similarity 79.1%; Pred. No. 5.9e-294;
Matches 2129; Conservative 0; Mismatches 445; Indels 116; Gaps 13;
QY 90 CGCGCCGACGCGCCGCGCATGGGGCTGAGGGCCGCCAGAGAGCGCTGTTCCGCGTG 149
Db 41 CGGAGGCCACCGCCCGCCGCGCATGGGGCTGAGGGCCGCCAGAGAGCGCTGTTCCCTCTG 100
QY 150 CGCTCATCGACGACGCTGGTGGCCCTGTTGCTGCCGAGCTGGCGAGAGAGCGCGAC 209
Db 101 CGCTCATCGACGACGCTGGTGGCCCTGTTGCTGCAGAGCTGGCGCGAGAGAGCGCTGAC 160
QY 210 CTGGTCTCTTTCTTCTTGGTGGCTTCTGTTGAGCATTTTCTGGTGTCAACCGCGTC 269
Db 161 CTGGTCTCTTCTTCTTGGTGGCTTCTGTTGAGCATTTTCTGGTGTCAACCGCGTC 220
QY 270 ATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGAGCCCGCCCGCCCGCTGGC 329
Db 221 ATCCCGACCAACGTTCCCGAGCTCACCTTCCAGCCCGAGCCCGCCCGCTGGT 280
QY 330 GGCTCACCTACTTCTCCGTCGGCGACCTGTCTATCATCGCCCGCTCTATATGCCGCTTC 389
Db 281 GGCTCACCTACTTCTCCGTCGGCGACCTGTCTATCATCGCCCGCTCTATATGCCGCTTC 340
QY 390 ACCGCCAGATCCGAGCGCCGCTCGACCTGTCTCTCTATCTCGAGAAGGGGTGTCTCC 449
Db 341 ACCGCTCAGATCCGCGCGCTGTGGACCTCTCTCTCTATCTCGAGAGGAGGTGTTCT 400
QY 450 ACCGCTGAGCTGGTGAAGAGGCTCTCCGATGTCTATATGGAACAGCTCAGCCCTCTCTAC 509
Db 401 AGTCGGAACCTGTTAAAAAAGGTCTCGGATGTCTATATGGAACAGCTCAGCCCTCTCTAC 460

Db 1477 AGGAGATTCTAAGGAATCTTTTGAAGTGCCCAATGACGTATCCCCCAACCTGCTGAAGG 1536
Qy 1243 AGGCAGCCAGCTCTGAGCGCGGAGGAGCGCGGGGAGCAAGACGAGGCAACC 1302
Db 1537 AGGCGGACCGCTGCTAGAACAGCGCAAGAGCGGAGCTGGGGAGCAAGCCAGGGCAACC 1596
Qy 1303 AGAGCCAAAGGTTCCGCCCTCCAGGACCCCTGAGTGTCTTTCGCCCACTGCTCGATTCTACG 1362
Db 1597 AGGSCAAAGGTTCTGCTCTCCAGGACCCAGAAATGCTTTTGGCCCACTGCTCGATTCTATG 1656
Qy 1363 AGGCATCTGCAATGAGGAGGAGGAGTCCAGGCGCTGTGTCAGCTGGCTGGGCCA 1422
Db 1657 ATGSCATCTGCAATGAGGAGGAGGTAGCCCCACACAGTACTGCAGCTAGGCTGGGCAA 1716
Qy 1423 CTTTCTTGTGTCAGTCTCCAGGCGTTTTCAGGGACAGGTGCGGCAGAGGTGCGCATAG 1482
Db 1717 CTTTCTTGTGTCAGTCTCCAGGCGCTTCAGGGACAGGTGCGGCAGAGGTGCGCATAG 1776
Qy 1483 TGAGCCGAGAGCGGAGCGGCCGAGGCGGAGGCGGTCGGGGAGGAAGCCCGGGAAG 1542
Db 1777 TTAGCCGGGAGACAGAGCGGCCGAGGCTGAAGAACCATGGGGGATGAAGCCCGAGAAG 1836
Qy 1543 GCCGGCGGGGGCCACAGCGGGGAGTCCAAAGCCAGAGGAGCCCGCCGCCCAAGAGC 1602
Db 1837 GCCGTGCGGCTGTGTCGCGGAGAGAGTCCAAAGCCTGAGGAGCCACACACCGCAAGAGC 1896
Qy 1603 CAGCACTGGACAAGGGCTTGGGCACCGCCAGGCTGAGGTGTCAGGACCCCGCCGGAAGC 1662
Db 1897 CTGCTTGGACAGGGCCCGGCTCAGACACAGTGCAGGGTGGGACCACTAGGAAAA 1956
Qy 1663 CTCCTGGGACTGTGCTGGCACAGCCC --- GAGGCCCTTGAAGTGGGACGACGCTCAGG 1719
Db 1957 CGTCAGGAGTGTCCCAGGTACTACCCGCGGTGCGCCAGGAAGTCCGCAATGCTGCTCAGG 2016
Qy 1720 TGCAGACACCGCAGCATCACCACCGCGGAGGCTCAGTCTCAGTTCAGAGTGA 1779
Db 2017 CTCAGACACTTGCAGCATCGCCACCCCGAGGCGCCAGTCTCAGTTCAGAGTGA 2076
Qy 1780 AGATGAAGGATGAAGGAGTCTGTGTGGCCACCAAGATCAACTGAGGCGCCATCAAGC 1839
Db 2077 AGATGAAGGATGAAGGAGTCTGTGTGGCCACCAAGATCAACTGAGGCGCCATCAAGC 2136
Qy 1840 TGCAACTCAGCGCAGTTCGCAAGTGCAGATGAAGAGCAGAAAGTGCACCCCTAGTG 1899
Db 2137 TGCAACTCAGCGCAGTTCGCAAGTGCAGATGAAGAGCAGAAAGTGCACCCACCGAGC 2196
Qy 1900 ACTACACTCTGCTTCTCCTCAAGCGGAGCGCAAGGCTCTGAACCTCTGAGGACTTCG 1959
Db 2197 ACTACACTCTGCTTCTCCTCAAGCGGAGCGCAAGGCTCTGAACCTCTGAGGACTTCG 2246
Qy 1960 GACCGCTTGTGGGACCGAGGCTCGGCTTAGTCTCCCACTGAGCGCCATGTTCTGCC 2019
Db 2247 ATTCTGTGCAAGGCCCGAGGCTGCGCTCAGTCTCCAGCGGCTTACCCCTCATCCTA 2306
Qy 2020 CCCAGCCCAAGGGGACAGGCGCTCAGCTCTACCCAAACCTTAGTTCGCGTCCCGAGTA 2079
Db 2307 ATTCAAGCCTAGAC -TCGCGCTCAGTCTTAGGGGACCGAGGCTGCTCAGTGA 2365
Qy 2080 CAGTCTGTATCAACCCACAGATTTTCTCAGCTCAGAACCCAGGCTGCTGCCAGCTGT 2139
Db 2366 G-----GAGCTTCAATTTAGCCCTCTTTCAGGCGCCAGAGGCTGCTCAGTCT 2415
Qy 2140 TAGAATATAGTCTCTTCTCCAGAAATCCAGCGGGGCAATGGAACCTCAGCTGGGTC 2199
Db 2416 TAGAATATAGTCTCTTCTCCAGAAATCCAGCGGGGCAATGGAACCTCAGCTGGGTC 2459
Qy 2200 CTAATTACCACTCTTTAAAGGCCCGAGGCTTAGAACCCAGGCTCTCTCGGAACCGCT 2259
Db 2470 CAATCCCCCGAAATTCAGTGGACCATGAC --- AGGAGCCCAAGTCAAGGCTGAGGTTG 2527
Qy 2260 CAGCTAGAGCCGACCAACGTTACTCAGGCGCTCTCCAGGCTTGTAGGAGCTGAGGTTTC 2319
Db 2528 GCATTCCTCCCACTTGGGGCTAGATTGACCTTAGGCC----- 2565

Qy 2320 ACCCTTAACCAAGGAGGACACAGTCCCACTCCAGCCCGGAGGCTTAGGACCACTTCAG 2379
Db 2566 -----CAGAGCACAGTTGCTACT-----GGAACCAAGGACCACTTCAG 2604
Qy 2380 CCCTAGGAGTATATTTCCGCACTT--CAGAAATTCATATCTTGGGANTCAAGCTCCCT 2437
Db 2605 TCCTAGGGGTAGATTTTCCGCACTTTATAGAGTTCCATCTTTTGGGAATCAAAATCCCTT 2664
Qy 2438 GCCCAAAATCACTCAGTCTGCTTCCAGAAATTTGGAAATCTTCTCTCTCTCTCTCT 2497
Db 2665 ACCCAAGAACCTTACATATTTGCTGCGGAAATTTGGAATC-----TTCCTCTCTCTCTCT 2719
Qy 2498 ATCCGAGTCTGGGACACAAACTCCGCCCGGAGCTTATGAGCATCTGAGCCCGGCT 2557
Db 2720 AGCCCTGAGTACAGGACATGAAGCTCCGACTCTAATCTCTGAATCTGAGCCCTCGGCC 2779
Qy 2558 CTCCTGACAAACTGGCCCGGATCAGAGGAGACCTCCCTCCGACCTCTTGGGAAAC 2617
Db 2780 CTCCTGACAAACTGGCGCAAGAACAGAGAGGACCT--CTTCCAGCCCTCTGGGAAC 2837
Qy 2618 TCCAGAGTCCAGCCCATCTCGGAGCATCCCGGAGGAAATCTGCAGAGGGTTAGGAGT 2677
Db 2838 TCCTGAGGTC-----CCAGAGGAAATCTGCAGCGGTTGGGGAG- 2878
Qy 2678 GGTGACAAGAGCCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2737
Db 2879 -----GGCCAGCCT 2927
Qy 2738 AGATGAATACATTTTGTGA 2757
Db 2928 AGATAAATGCAATTTTGTGA 2947

RESULT 10

AF072755 2577 bp mRNA ROD 11-JUL-1998

LOCUS Mus musculus MENIN (Men1) mRNA, complete cds.

DEFINITION AF072755

ACCESSION AF072755

VERSION AF072755.1 GI:3309058

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 2577)

Bassett,J.H.D. and Thakker,R.V.

cDNA sequence of mouse MEN1 gene

Unpublished

2 (bases 1 to 2577)

Bassett,J.H.D. and Thakker,R.V.

Direct Submission

Submitted (15-JUN-1998) MRC Molecular Endocrinology, Imperial

College School of Medicine, Hammersmith Hospital DuCane Road,

London, NA W12 0NN, United Kingdom

Location/Qualifiers

1. .2577

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="19"

/notes="map location based on human syntony"

1. .2577

/gene="Men1"

46. .1881

/gene="Men1"

/notes="similar to Homo sapiens MENIN"

/codon_start=1

/product="MENIN"

/protein_id="AAC26001.1"

/db_xref="GI:3309058"

/translation="MGLKAAQKTLFPLRSIDDDVVRFLPAELGHEEPDI.V.I.SLVIGPV

EHLFVNRVITNPVLPFQPPGLTFFPVADLSITAAALYARPTAQRGAVD

LSLYPREGGVSSRELKVKVDIWNLSLSRSFYFKDRAIIQSLFSFITCTKLDSSGVAFA


```

Db 1935 CCCTCATCTCTTAATCAAGCCTAGAC-TCCGCCCTCAGTCTCAGGAGGCCAGAGCTGT 1993
QY 2067 CCGTCCCGAGTACAGTCTGTATCAACCCACAGATTTTCTCCAGCTCAGAACCCAGGGCT 2126
Db 1994 CTGACAGTACTAG-----GAGCCTTCAATTTAGCCCTCTTACAGGCCAGAGCC 2043
QY 2127 CTGCCCGAGTCTAGATATAGTCTCTTCTCCAGAAATCCAGCGGGCCAAATGGAAAC 2186
Db 2044 CTGGCTCAGTCTTAGAATGTAGTGTATCCCA-----CCAGCGCGCAATTGGAAC 2097
QY 2187 CTCAGCTGGTCTCTTAATACAGTCTTTAAAGCCAGCCCTCAGAACCCAGCTCT 2246
Db 2098 CTATCTAGTCTCAATCCCGCAATTTAGTGACCATGAC--AGGAGCCAGTCAAGC 2155
QY 2247 CTTCCGNAACGCTCACCTAGAGCCAGACCAACGTTACTCAGGGCTCTCCACGCTGTAG 2306
Db 2156 CCT-----GACAGTATGGCA 2171
QY 2307 GAGCTGAGGTTTACCCTTAACCCAGGAGCACAGTCTCCACCTCAGCCCGGGAGCC 2366
Db 2172 TTCTCCACCTTGGGGCTAGATTGACCTTAGCCAGACACAGTTGCTACCTGGAAAC 2231
QY 2367 TAGACACACTCAGCCCTAGAGTATATTTCCGGACTT--CAGAAATCCATATCTCCGA 2424
Db 2232 AAGACCAATTCAGTCCCTAGGGTAGATTTCCGCACTTTATAGAGTTCCATCTCTTGGGA 2291
QY 2425 ATCAAGCTCCCTGCCCAATAACTTCACTCTCTCTCAGAAATTTGGAATCTTAGTT 2484
Db 2292 ATCCAAATCTTACCCCAAGAACTTACATATTTGCTGCCGAAATTTGGAATC-----T 2346
QY 2485 TCCTCTCTCTTCTGATCCCGAGTCTGGGACACAAACTCCGCCCCAGCCCTATGAGCATCC 2544
Db 2347 TCCTCTCTCTCCTAGCCCTGAGTACAGGACATGAAGCTCCGACTCTAATCTGTGAATAC 2406
QY 2545 TGACCCCGCCCTCTCTCAGCAAACTGCCCGGATCAGACAGGACCTCCCTCCGA 2604
Db 2407 TGAGCTCGCGCCCTCTCTGACCAAACTGGCGCAAGAACAGACAGGAGGACCT--CTTCAG 2464
QY 2605 CCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGAGCATCCCGGAGGAAATCTGCAG 2664
Db 2465 CCCTCTGGGAACCTCTCTGAGTC-----CCAGGAGGAAATCTGCAC 2506
QY 2665 AGGGTTAGAGTGGGTGACAGAGCTGTATCTCTCTCTTTTGTACATAGATTTATTT 2724
Db 2507 GGGGTGGGGAG-----GGCCAGCCTCTCTCTCTTTTGTACATAGATTTATTT 2554
QY 2725 TTCAAGTCCCAAGAA 2739
Db 2555 TTCAGTCCCAAGAA 2569

RESULT 11
AF130369
LOCUS AF130369 2261 bp mRNA ROD 07-OCT-1999
DEFINITION Rattus norvegicus strain Wistar MEN1 tumor suppressor (Men1) mRNA,
alternative splice product, complete cds.
ACCESSION AF130369
VERSION AF130369.1 GI:6014460
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2261)
AUTHORS Karges,W., Maier,S., Wissmann,A., Dralle,H., Dosch,H.M. and
Boehm,B.O.
TITLE Primary structure, gene expression and chromosomal mapping of
rodent homologs of the MEN1 tumor suppressor gene
JOURNAL Blochim. Biophys. Acta 1446 (3), 286-294 (1999)
MEDLINE 99453728
REFERENCE 2 (bases 1 to 2261)
AUTHORS Karges,W., Maier,S., Wissmann,A., Dralle,H., Dosch,H.M. and

```

```

Boehm,B.O.
Direct Submission
Submitted (24-FEB-1999) Division of Endocrinology, University of
Ulm, Robert Koch Strasse 8, Ulm 89081, Germany
FEATURES
Location/Qualifiers
1..2261
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
1..2261
/gene="Men1"
CDS
378..2210
/gene="Men1"
/note="MEN1; alternatively spliced"
/codon_start=1
/product="MEN1 tumor suppressor"
/protein_id="AA01353.1"
/db_xref="GI:6014461"
/translation="MGLKAAQTLFPLRSIDDVRLFAAEIAGKEHMDLVLLSLVLFV
EHFLAVNRVPTNPBELTQPPAPPPGLTVFPVADLSIIAALYARFTAGTARGAVD
LSLYPREGVSSRELKVKVSDVWNSLSYKDRAHIOSLFSFTGTGKLDSSGVAF
VVGCOALGLRDVHLALSDHAWVVGSGEOTAEVTHGKQNEDRGOTVAGVAER
SHLYLKGSMRCORKEVAFVCAINPSIDLTLDSLELLQLOKLLWLYDLGLERY
PWALGNLADLEELPTGPRDPLTLHKGIASAKTYQDEHIYPYMLAGYHCRNRV
REALQAWADTATVIOYNYCDEDEEYKEFFVANDVIPNLKKAASLLAEGERPGE
QAQGTGGGSAQDPECFHLLRFYDGIKWBEGSPVPLHVGWATFLVQSILGRFEG
VQKVIHVSREAAEAEBPWGDEARERRRPRRESKPEEPKPKKPAALDKPGSGQ
SAGSPPRKTSVTSCTARGTEVSSAAQAPAAAPPEPVLITFOSERMKMKMKELV
ATKINSALKQLTQSGVQMKQKVSFSDYTLFLKQRKQL"
BASE COUNT 479 a 676 c 689 g 417 t
ORIGIN
Query Match 54.9%; Score 1522; DB 94; Length 2261;
Best Local Similarity 87.2%; Pred. No. 2.8e-276; Indels 0; Gaps 0;
Matches 1669; Conservative 0; Mismatches 245;
QY 43 GCAGCGCGCCCGCCACCTATTTCCAGGCTCTCGGGGCGAGGGCGCCGCGCCACGCG 102
Db 310 GGAGGGGCTCTCTCGCGGAGCGCGACCTCTCTCGGCTTCCCGCAGCGCCACGCG 369
QY 103 CCGCGCGCATGGGGCTGAAGCGCGCCAGAGAGCTGTTCCCGCTGGCTCCATCGACG 162
Db 370 CCGCGCGCATGGGGCTGAAGCGCGCCAGAGAGCTGTTCCCTCTCGGCTCTATCGACG 429
QY 163 ACCTGTGTGCGCTGTTGCTGCGAGCTGGCGGCGAGAGAGCGGACCTGCTGCTCTT 222
Db 430 ACCTGTGTGCGCTGTTGCTGCGAGCTGGCGGAGCTGGCGGAGAGAGCTGCTGCTCT 489
QY 223 CCTTGTGTGCGCTGCTGCTGAGCATTTTCTGGCTGTCAACCGGCTCATCCCTACCAACG 282
Db 490 CCTTGTGTGCGCTGCTGAGCATTTTCTGTGCTGTCAACCGCTGTCTATCCCAACG 549
QY 283 TTCCGAGCTCACTTCCAGCGCCCGCCCGCCCGCCCGCTGGCGGCTCTACCTACT 342
Db 550 TGGCGGAGCTCACTTCCAGCGCCCGCCCGCCCGCCCGCTGGCTGGCTCTACCTACT 609
QY 343 TTCCGCTGGCGGACCTGCTCTATCATCTCCCGCTCTATGCTCCGCTTCCAGCGCCAGATCC 402
Db 610 TCCAGTGGCGGACCTATCCATCATCTGCTGCTCTATGCTCCGCTCATCCGCTCAGATCC 669
QY 403 GAGCGCGCTGCGACCTGCTCTCTATCTATCTCGAGAGGGGGTGTCTCCAGCGGTGAGCTGG 462
Db 670 GCGGTGTGTGGACCTCTCTCTCTATCTCGAGAGGGGGGCTTTCAGTTCGCGCACTGG 729
QY 463 TGAGAGAGTCTCCGATGTCATATGGAACAGCTCACCGCTCCCGCTCTCAAGTACGG 522
Db 730 TAAAGAGTCTCGGATGTATATGGAACAGCTCACCGCTCTCTACTTCAAGAGCGCGG 789
QY 523 CCCACATCCAGTCTCTCTTCAAGTTCATCAGAGCAGCAAAATTCAGACAGTCCCGGTGG 582
Db 790 CCCACATCCAGTCTCTCTTCAAGTTCATCAGAGCAGCAAAATTCAGACAGTTCGCGCGCTGG 849

```


ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 1906)
AUTHORS Karges,W., Maier,S., Wissmann,A., Dralle,H., Dosch,H.M. and
Boehm,B.O.

TITLE Primary structure, gene expression and chromosomal mapping of
rodent homologs of the MEN1 tumor suppressor gene

JOURNAL Biochim. Biophys. Acta 1446 (3), 286-294 (1999)

MEDLINE 99453728

REFERENCE 2 (bases 1 to 1906)
AUTHORS Karges,W., Maier,S., Wissmann,A., Dralle,H., Dosch,H.M. and
Boehm,B.O.

TITLE Direct Submission

JOURNAL Submitted (24-FEB-1999) Division of Endocrinology, University of
Ulm, Robert Koch Strasse-8, Ulm 89081, Germany

FEATURES
source
1..1906
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="RIN 5AH"
/cell_type="islet"
1..1906
/gene="Men1"
23..1855
/gene="Men1"
/note="MEN1: alternatively spliced"
/codon_start=1
/product="MEN1 tumor suppressor"
/protein_id="AA01354.1"
/db_xref="GI:6014463"
/translation="MGLKAAQKTFPLRSIDVVVFLFAELGRPEPDLVLLSLVLGFV
EHLAVNRVITPNVPELTFQSPAPDPPGLTYEPVADLSIIAALYAFQIRGAVD
LSLIPREGVSSRELKVKVSDVINSLRSYFKDRAHQSLFSFTIGKLDSSGVAF
WVLQAGLGRDHLALSEDHAWVYFNGSEQTAEVTHWKGNDNRGQTVNAGVAER
SWYLGKSYMDRMEYAFVCAINPSIDILTDSLELLQLOKLLWLLYDLHLERY
PMALGNLADLELPDTPGRDPLTLYHKGIAAKTYQDEHYIPYMLAGYHCRNRV
REALQWADTATVODNYCRDEDEIYKEFEVANDVIPNLLKEASLLEAGEPRPE
QAQGTGGGSAQLOPECFALLRFYDGKWEESPTPLVHGNATFLVQSLGRFEGQ
VRQKHVYSREAEAEPEPWDEAREGRRRPREPPEPPPKKPAKJGPGSGQ
SAGSPPKRTVSGTARGTEVSAQAQAPAPASPPPEPGLVTFQSEKMGKMLLY
ATKINSIAIKLQLTAQSOVMKQKVSFPTDYSFLKRQRKL"

BASE COUNT 419 a 582 c 552 g 353 t

ORIGIN

Query Match 54.8%; Score 1518.2; DB 94; Length 1906;
Best Local Similarity 88.3%; Pred. No. 1.5e-275;
Matches 1649; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 90 CGCGCCGACCGCGCGCCATGGGCTGAAGCGCGCCGACAGACGCTGTTCCGCGTG 149
DB 2 CGAGGCGACCGCGCGCCGCGCCATGGGCTGAAGCGCGCCGACAGACGCTGTTCCCTCG 61

QY 150 CGCTCCATCGACGAGTGTGGCGCTGTTTGTCTGCGAGCTGGCGGAGAGAGCGCGAC 209
DB 62 CGCTCTATCGACGAGTGTGGCGCTGTTGCTGCGAGCTGGCGGAGAGAGCGCTGAC 121

QY 210 CTGGTGCTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
DB 122 CTGGTGCTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181

QY 270 ATCCCTPACCAAGCTTCCGAGCTCACTTCCAGCCGCGCGCGCGCGCGCGCGCGCG 329
DB 182 ATCCCGACCAAGCTTCCGAGCTCACTTCCAGCCGCGCGCGCGCGCGCGCGCGCG 241

QY 330 GGCCTCACCTACTTTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 389
DB 242 GGCCTCACCTACTTTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301

QY 390 ACGGCCAGATCGGAGCG 449
DB 1382 AAGGTGCATAGTTAGCCGGGAAGCGGAGCGAGCGGAGCGGAGCGGAGCGGAGCGG 1441

DB 302 ACCGCTCAGATCCGGGTGCTGTGGACCTCTCCCTCTATCTCTCGAGAGGGGGCGTTTCC 361
QY 450 AGCCGTGAGCTGGTGAAGAAGGTCTCCGATGTATATGAACAGCCTCAGCCGCTCTAC 509
DB 362 AGTCGGAACTGGTAAAAAGGTCTCGGATGTATATGAACAGCCTCAGCCGCTCTAC 421
QY 510 TTCAAGGATCGGGCCACATCCAGTCCCTCTTTCAGTTCATCACAGCCACCAATTTGGAC 569
DB 422 TTCAAGGACCGGGCCACATCCAGTCTCTTTCAGTTCATCACAGCCACTAAACTGGAC 481
QY 570 AGCTCGGTGTGGCTTTCGTGTGGTGGGCTGCCAGGCCCTGGGTCTCGGGATGTC 629
DB 482 AGTTCGGGCTGGCTTTCAGTGGTAGGGCTGCCAGGCTCTAGGTCTCAGAGATGC 541
QY 630 CACCTCGCCCTGTCTCAGGATCATGCTGGGTAGTCTTTGGGCCCAATGGGAGCAGACA 689
DB 542 CATCTGGCCCTGTCCGAGGATCATGCTGGGTGGTGTCTGGGTCCCAATGGGAGCAGACG 601
QY 690 GCTGAGGTCACTTGGCAGCGCAAGGCAACGAGGACCGCAGGGGCCAGACAGTCAATGCC 749
DB 602 GCTGAGGTGACCTGGCATGGCAAAAGCAACGAGGACCGCAGAGGCCAGACAGTCAACGCT 661
QY 750 GGTGTGGCTGAGGGAGCTGGCTGTACCTGAAAGGATCATATGCGCTGTGACCGCAAG 809
DB 662 GGTGTGGCTGAGGGAGCTGGCTGTACCTGAAAGGATCGTACATGGCGCTGTGACCGCAAG 721
QY 810 ATGGAGGTGGCTTTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 869
DB 722 ATGGAGGTGGCTTTCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
QY 870 CTGGAGCTTCTGCACTGCGCAGCAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 929
DB 782 TTGGAACTGTGCAACTACAGCAGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
QY 930 GAAAGTACCCATGGCTTAGGAACTTGGGAACTTGGGAACTTGGGAACTTGGGAACTTGG 989
DB 842 GAAAGTACCCATGGCTTAGGAACTTGGGAACTTGGGAACTTGGGAACTTGGGAACTTGG 901
QY 990 GCGCGCCAGACCCACTCACTTACCACAAAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1049
DB 902 GGTGCGCCAGACCCGCTCACTTATCACAGGGGATGCTGCTGCTGCTGCTGCTGCTGCTG 961
QY 1050 CGGGATGAACACATCTACCCCTACATGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1109
DB 962 CAGGATGAACACATCTACCCCTACATGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1021
QY 1110 GTGCGGAAGCCCTGCGAGCTGGGCGGACCGGCCACTGTCTCAGGACTACAACACTAC 1169
DB 1022 GTGCGTGAAGCCCTGCGAGCCCTGGGCGGACACTGCCACTGTTATCCAAAGACTACAAC 1081
QY 1170 TGCGGGAAGACGAGGAGATCTACAAGGAGTCTTTTGAAGTAGCCCAATGATGTATCCCC 1229
DB 1082 TGTGCGGAGATGAAGAGATCTACAAGGATCTTTTGAAGTAGCCCAATGATGTATCCCC 1141
QY 1230 AACCTGCTGAAGGAGGACGACAGCTTGTGAGGCGGCGGAGGAGCGCGCGGGGAGCAA 1289
DB 1142 AACCTGCTGAAGGAGGACGACAGCTTGTGAGGCGGCGGAGGAGCGCGCGGGGAGCAA 1201
QY 1290 AGCCAGGCGACCCAGAGCCAAAGTTCCGCCCTCCAGGACCTCAGTGTCTGCGGCCACTG 1349
DB 1202 GCGCGGCGACCCAAAGGTTCTGCCCTCCAGGACCCAGAGTGTCTGCGGCCACTG 1261
QY 1350 CTGCGATTCTACGACGCGATCTGCAAAATGGGAGGAGGAGTCCCAAGCTGTGCTGTCAC 1409
DB 1262 CTGCGATTCTATGATGGCATCTGCAAAATGGGAGAGGATAGCCCCACACAGTACTGTCAC 1321
QY 1410 GTGGGTGGGCCACTTCTTGTGTCAGTCCCTAGGCGCTTTTTCAGGAGACAGGTGCGGCGAG 1469
DB 1322 GTGGGTGGGCAACTTCTTGTGTCAGTCCCTAGGCGCTTTTTCAGGAGACAGGTGCGGCGAG 1381
QY 1470 AAGGTGCGCATAGTGAAGCGAGGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGGAG 1529
DB 1382 AAGGTGCATAGTTAGCCGGGAAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 1441

```

QY 1530 GAAGCCCGGAAGCCGCGCGCGGCCCGCCACAGCGGGAGTCCAAGCCAGGAGGCCCGC 1589
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1442 GAAGCCCGGAAGGTCTCGCGCGGGTCCCGCGCGAGTCCAAGCTGAGGAGCCACCA 1501
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1590 CCGCCCAAGAACCCAGCACTGGACAAGGGCTGGGCACCGCCAGGCTGAGTGTACAG 1649
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1502 CCACCAAGAAGCCTGCGTTGGACAAGGGCCAGGCTCAGGGCAAGTGCAGGTCGGGA 1561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1650 CCCCCTCGAAGCCTCTCTGGGACTGCTCGTGGCACAGCCCGAGGCCCTGAAGTGGCAGC 1709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1562 CCTCTAGGAAAACGTCAGGACTGCTCTCAGTACTCCCGTGGCCACAGAGTCAGCAGT 1621
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1710 AGGCTCAGGTGCCAGCACCCCGCAGCATCACACCGCGGAGGCTCCAGTGCCTCCTTC 1769
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1622 GCTGCTCAGGCTCCAGCACCTGCAGCATCGCCACCCCGAGAGGCCCTGCCTCCTTC 1681
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1770 CAGAGTGAAGATGAAGGCATGAAGAGTGTGTGTGGCCACCAAGATCAACTCGAGC 1829
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1682 CAGAGTGAAGAAATGAAGGAATGAAGGAGCTACTGTGTGGCCACCAAGATCAACTCTAGT 1741
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1830 GCCATCAAGCTGCAACTCAGCGCACGTGCGAAGTGCAGATGAAGAGCAGAAAGTGTCC 1889
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1742 GCCATCAAGCTGCAACTCAGCGCACGTGCGAAGTGCAGATGAAGAGCAGAAAGTGTCC 1801
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1890 ACCCTAGTCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1949
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1802 ACACCCAGGACTACACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1861
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 1950 GGGGACT 1956
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1862 CGGGATT 1868
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

RESULT 14
HSU93237
LOCUS HSU93237 9180 bp DNA PRI 19-APR-1997
DEFINITION Human menin (MEN1) gene, complete cds.
ACCESSION U93237
VERSION U93237.1 GI:1945388
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9180)
AUTHORS Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufermi,S.-E.,
Collins,F.S., Emmert-Buck,M.R., Debelenko,L.V., Zhuang,Z.,
Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A.,
Weismann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S.,
Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.
TITLE Positional cloning of the gene for multiple endocrine
neoplasia-type 1
JOURNAL Science 276 (5311), 404-407 (1997)
MEDLINE 97258940
REFERENCE 2 (bases 1 to 9180)
AUTHORS Collins,F.S.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1997) National Human Genome Research Institute,
Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD
20892, USA

FEATURES
source
1. .9180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
repeat_region complement(1..68)
/rpt_type-dispersed
/rpt_family="ALU"
gene 1. .9180
/genes="MEN1"
repeat_region complement(85..168)
/rpt_type-dispersed
/rpt_family="MER42"
392..1285
/rpt_type-dispersed
/rpt_family="ALU"
complement(876..1004)
/rpt_type-dispersed
/rpt_family="ALU"
1680..1766
/genes="MEN1"
/number=1
join(1680..1766,2265..2747,4297..4505,4716..4844,
5177..5217,5298..5385,6025..6161,6623..6758,7196..7360,
7578..8874)
/genes="MEN1"
/notes="alternatively spliced"
join(1680..1766,2265..2732,4297..4505,4716..4844,
5177..5217,5298..5385,6025..6161,6623..6758,7196..7360,
7578..8874)
/genes="MEN1"
2265..2732
/genes="MEN1"
/number=2
2265..2747
/genes="MEN1"
/notes="alternate exon 2"
join(2288..2732,4297..4505,4716..4844,5177..5217,
5298..5385,6025..6161,6623..6758,7196..7360,7578..8060)
/genes="MEN1"
/codon_start=1
/product="menin"
/protein_id="AAC51229.1"
/db_xref="GI:1945389"
/translation="MGLKAAOKTLPRLSIDDVVRIFAELGREPDVLVLSLVLGFEV
EHFLAVNRVPTNPVPELTQPSADPPGGLTYFPVADLSITAAVARTAOIRGAVD
LSLIPREGVSSRELKVKVSDVIWNSLSRYFKDRAHIOQLSFTGCTKLDSSGAF
VVGACALGRDVLHLSDEHAWVFGNGEOTAEVTHGKNGEDRRGQTVNAGVAER
SMYLGSYMRCDRKMEVAFYCAINPSIDLHTDSLELLQLOQLLWLLYDLGLERY
PNALGNLADLELEPTPGRPDLTYLHKGIASAKTYRDEHIYPYMYLAGYHCRNRV
REALQAWADTATVODNYCREDEITYKEFFEVDVDPVLLKKAASLLEAGERPGE
QSQGTQSGSALQDPECFALHLLREYDGIKWEEGSPTVLHVGMATFLVQSLGRFEG
VQKRVIRVSEAAEAEPWEEAREGRRRPRRESKPEEPKPPKALDKGLGTGQ
GAVSGPPRPKPTAGTARGEGGTAQVPAAPSPPEPGLTTPQSEKMKMKELLY
ATKINSSAIKQLTAQSQVMKQKQVSTPSTYLSFLKQRKGL"
join(2288..2747,4297..4505,4716..4844,5177..5217,
5298..5385,6025..6161,6623..6758,7196..7360,7578..8060)
/genes="MEN1"
/notes="alternatively spliced"
/codon_start=1
/product="menin"
/protein_id="AAC51230.1"
/db_xref="GI:1945390"
/translation="MGLKAAOKTLPRLSIDDVVRIFAELGREPDVLVLSLVLGFEV
EHFLAVNRVPTNPVPELTQPSADPPGGLTYFPVADLSITAAVARTAOIRGAVD
LSLIPREGVSSRELKVKVSDVIWNSLSRYFKDRAHIOQLSFTGCTKLDSSGAF
VVGACALGRDVLHLSDEHAWVFGNGEOTAEVTHGKNGEDRRGQTVNAGVAER
SMYLGSYMRCDRKMEVAFYCAINPSIDLHTDSLELLQLOQLLWLLYDLGLERY
PNALGNLADLELEPTPGRPDLTYLHKGIASAKTYRDEHIYPYMYLAGYHCRNRV
REALQAWADTATVODNYCREDEITYKEFFEVDVDPVLLKKAASLLEAGERPGE
QSQGTQSGSALQDPECFALHLLREYDGIKWEEGSPTVLHVGMATFLVQSLGRFEG
VQKRVIRVSEAAEAEPWEEAREGRRRPRRESKPEEPKPPKALDKGLGTGQ
GAVSGPPRPKPTAGTARGEGGTAQVPAAPSPPEPGLTTPQSEKMKMKELLY
ATKINSSAIKQLTAQSQVMKQKQVSTPSTYLSFLKQRKGL"
variation
2352
/genes="MEN1"
/standard_name="L22R"
/notes="Lys22Arg"
/replace="g"
2534..2537
/genes="MEN1"
/standard_name="357del4"
/notes="4 bp deletion"
/replace="n"

```

```
variation 2593
/gene="MEN1"
/standard_name="416delC"
/note="1 bp deletion"
/replacement=""
2641..2643
/gene="MEN1"
/standard_name="K119del"
/note="3 bp deletion"
/replacement=""
2689
/gene="MEN1"
/standard_name="512delC"
/note="1 bp deletion"
/replacement=""
complement(3117..3725)
/rpt_type="dispersed"
/rpt_family="ALU"
4297..4505
/gene="MEN1"
/number=3
4445
/gene="MEN1"
/standard_name="W198X"
/standard_name="Trp198stop"
/replacement=""
4476..4479
/gene="MEN1"
/standard_name="735del14"
/note="4 bp deletion"
/replacement=""
4716..4844
/gene="MEN1"
/number=4
5177..5217
/gene="MEN1"
/number=5
5298..5385
/gene="MEN1"
/number=6
6025..6161
/gene="MEN1"
/number=7
6134
/gene="MEN1"
/standard_name="1132delG"
/note="1 bp deletion"
/replacement=""
6623..6758
/gene="MEN1"
/number=8
6660..6662
/gene="MEN1"
/standard_name="E363del"
/note="3 bp deletion"
/replacement=""
7196..7360
/gene="MEN1"
/number=9
7316
/gene="MEN1"
/standard_name="W436R"
/note="Trp436Arg"
/replacement=""
7317
/gene="MEN1"
/standard_name="W436X"
/note="Trp436stop"
/replacement=""
7578..8874
/gene="MEN1"
/number=10
7806
variation
```

```
BASE COUNT 1936 a 2604 c 2612 g 2027 t 1 others
ORIGIN

Query Match
Best Local Similarity 47.3%; Score 1312.2; DB 97; Length 9180;
Matches 1341; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1382 GGAGGCGAGTCCACGCTGTGCTGACGTGGGCTGGGCCACCTTTCTTTGTGCAGTCCCT 1441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7499 GGATGCCAGAGCAGGGTCTGGAGTTCCAGCACCTGGCCGCGCAACCTTGTCTCACCTT 7558

QY 1442 AGGCCGTTTGGAGGACAGGTGCGGCAGAGGTGGCATAGTACGCCAGAGGCCGAGGC 1501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7559 GCTCTCCCACTGGCCAGGTGCGGAGAGGTGCGCATAGTACGCCAGAGGCCGAGGC 7618

QY 1502 GGCCGAGGCCGAGGAGCGTGGGCGAGGAAGCCCGGGAAGCCGCGGGGCCGAGC 1561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7619 GGCCGAGGCCGAGGAGCGTGGGCGAGGAAGCCCGGGAAGCCGCGGGGCCGAGC 7678

QY 1562 GCGGAGTCCAAAGCCAGAGGAGGCCCGCGCCCAAGAACCCAGCAGCTGCAAGGGCCT 1621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7679 GCGGAGTCCAAAGCCAGAGGAGGCCCGCGCCCAAGAACCCAGCAGCTGCAAGGGCCT 7738

QY 1622 GGGCACGCCGAGGTGTCAGTGTGAGGACCCCGGGAAGCCCTCTCTGGACTGTCTGCTGG 1681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7739 GGGCACGCCGAGGTGTCAGTGTGAGGACCCCGGGAAGCCCTCTCTGGACTGTCTGCTGG 7798

QY 1682 CACAGCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCCAGCACCCGCGAGCATCAC 1741
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7799 CACAGCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCCAGCACCCGCGAGCATCAC 7858

QY 1742 ACCGCCGAGGTCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCT 1801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7859 ACCGCCGAGGTCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGCGATGAAGGAGCT 7918

QY 1802 GCTGTGGCCACCAAGATCACTCGAGGCCCATCAAGTGAACCTCAGGCGACAGTCGCA 1861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7919 GCTGTGGCCACCAAGATCACTCGAGGCCCATCAAGTGAACCTCAGGCGACAGTCGCA 7978

QY 1862 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTCTCTCTCTCTCTCAA 1921
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7979 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTCTCTCTCTCTCAA 8038

QY 1922 GCGGAGCCGCAAGGCCCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 1981
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8039 GCGGAGCCGCAAGGCCCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 8098

QY 1982 TCCGCTTAGTCCCACTCTGAGCCCATGTCTGCCCCCAGCCCAAGGGGACAGGCC 2041
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8099 TCCGCTTAGTCCCACTCTGAGCCCATGTCTGCCCCCAGCCCAAGGGGACAGGCC 8158

QY 2042 TCACCTTACCAAAACCTTAGGTTCCTCGGTCCAGTACAGTCTGTATCAAAACCCAGCAT 2101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8159 TCACCTTACCAAAACCTTAGGTTCCTCGGTCCAGTACAGTCTGTATCAAAACCCAGCAT 8218

QY 2102 TTTCTCCAGCTCAGAAACCCAGGGCTCTGCCCCAGTCTGTAGTAATATAGTCTCTCTCC 2161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8219 TTTCTCCAGCTCAGAAACCCAGGGCTCTGCCCCAGTCTGTAGTAATATAGTCTCTCTCC 8278

QY 2162 AGATCCAGCGGGCCCAATGAAACCTCAGCTGGGTCTCTAATTACCAGTCTTTAAAGGC 2221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8279 AGATCCAGCGGGCCCAATGAAACCTCAGCTGGGTCTCTAATTACCAGTCTTTAAAGGC 8338

QY 2222 CCAGCCCTAGAAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 2281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8339 CCAGCCCTAGAAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 8398

QY 2282 ACTCAGGGCTCTCTCCAGCTTGTAGGAGTGTAGGTTTACCCCTTAACCCAGGAGCACA 2341
```

Db 8399 ACTCAGGGCTCTCCAGCTGTAGGAGCTGAGGTTTACCCCTTAACCCAGGAGCACA 8458
QY 2342 GGTCCACCTCCAGCCGGGAGGCTTAGGACCTAGCAGCCCTAGGAGTATATTTCCGCA 2401
Db 8459 GGTCCACCTCCAGCCGGGAGGCTTAGGACCTAGCAGCCCTAGGAGTATATTTCCGCA 8518
QY 2402 CTTTCAGAAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAATAACTTTCAGTCTGCT 2461
Db 8519 CTTTCAGAAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAATAACTTTCAGTCTGCT 8578
QY 2462 TCCAGAAATTTGGAATCCTAGTTCTCTCTTCGATATCCCGAGCTGGGACACAAACT 2521
Db 8579 TCCAGAAATTTGGAATCCTAGTTCTCTCTTCGATATCCCGAGCTGGGACACAAACT 8638
QY 2522 CCGCCCGCAGCTATGAGCTCTGAGCCCGCCCTCTCTCGCAAACTGCGCCCGGA 2581
Db 8639 CCGCCCGCAGCTATGAGCTCTGAGCCCGCCCTCTCTCGCAAACTGCGCCCGGA 8698
QY 2582 TCAGAGCAGGACCTCCCTCCGACCTCTCGGAACCTCCAGAGGTCAGCCCATCTCGG 2641
Db 8699 TCAGAGCAGGACCTCCCTCCGACCTCTCGGAACCTCCAGAGGTCAGCCCATCTCGG 8758
QY 2642 AGATCCCGAGGAATCTCAGAGGGTTAGGAGTGGTGACAAGAGCTGATCTCTTC 2701
Db 8759 AGATCCCGAGGAATCTCAGAGGGTTAGGAGTGGTGACAAGAGCTGATCTCTTC 8818
QY 2702 CTGTTTGTACATAGATTATTTTTCAGTTCCCAAGAAAGATGAATCATTTTGTAAAAA 2761
Db 8819 CTGTTTGTACATAGATTATTTTTCAGTTCCCAAGAAAGATGAATCATTTTGTAAAAA 8878
QY 2762 AAAAAAAA 2770
Db 8879 AATATAAA 8887

RESULT 15
AC000134/c
LOCUS AC000134 203300 bp DNA PRI 06-MAY-1999
DEFINITION Homo sapiens Chromosome 11q13 BAC Clone 137c7, complete sequence.
ACCESSION AC000134
VERSION AC000134.14 GI:4755212
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203300)
AUTHORS Crabtree, J. and Roe, B.A.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 203300)
AUTHORS Chandrasekharappa, S.C., Guru, S.C., Manickam, P., Olufemi, S.-E. and Collins, F.S.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 3 (bases 1 to 203300)
AUTHORS Emmert-Buck, M.R., Debelenko, L.V., Zhuang, Z., Lubensky, I.A. and Liotta, L.A.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 4 (bases 1 to 203300)
AUTHORS Weismann, J. and Boguski, M.S.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 5 (bases 1 to 203300)
AUTHORS Agarwal, S.K., Kester, M., Kim, Y.S., Heppner, C., Dong, Q., Spiegel, A.M., Burns, L.A. and Marx, S.J.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 6 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission

JOURNAL Submitted (12-MAR-1997) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 7 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 8 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (24-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 9 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 10 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 11 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 12 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 13 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On May 6, 1999 this sequence version replaced gi:4753208.
FEATURES
Source
1..203300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="137c7"
BASE COUNT 49152 a 57085 c 53982 g 43081 t
ORIGIN
Query Match 45.2%; Score 1253.2; DB 85; Length 203300;
Best Local Similarity 96.2%; Pred. No. 3e-226;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;
QY 1382 GGAGGGCAGTCCACGCCCTGCTGCTGCACGTGGGCTGGGCCACCTTTCTTGTGAGTCCCT 1441
Db 156588 GGATGGCCAGAGCAGGGTCTTGAGTTCCAGCCACTGGCGGCAACCTTGCTCTCACCTT 156529
QY 1442 AGCCCGTTTTGAGGGACAGGTGGCGGAGAGGTGCCATAGTGGCCGAGAGCCGAGGC 1501
Db 156528 GCTCTCCCACTGGCCAGGTGGCGGAGAGGTGCCATAGTGGCCGAGAGCCGAGGC 156469
QY 1502 GGCCGAGGCGGAGAGCCGTGGGGCGGAGGAAGCCCGGAGCGGCGGCGGCCACG 1561
Db 156468 GGCCGAGGCGGAGAGCGGTGGGGCGGAGGAAGCCCGGAGCGGCGGCGGCCACG 156409
QY 1562 GCGGGAGTCCAAAGCCAGAGGAGCCCCCGGCCCAAGAGCCACTGCACAGGCGCT 1621

Db 156408 GCGGGAGTCCAAAGCCAGAGAGCCCCCGCCGCGCCCAAGAAAGCCAGCACCTGGACAAGGGCCT 156349
QY 1622 GGGACCGGCCAGGGTGCAGTGTCAAGGACCCCGCCGGAAGCCTCTCGGGACTGTGCGTGG 1681
Db 156348 GGGACCGGCCAGGGTGCAGTGTCAAGGACCCCGCCGGAAGCCTCTCGGGACTGTGCGTGG 156289
QY 1682 CACAGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 1741
Db 156288 CACAGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 156229
QY 1742 ACCGCCGAGGGTCCAGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGGCATGAAGAGCT 1801
Db 156228 ACCGCCGAGGGTCCAGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGGCATGAAGAGCT 156169
QY 1802 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 1861
Db 156168 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 156109
QY 1862 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCTAGTGACTACACTCTCTTTTCTCTCAA 1921
Db 156108 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCTAGTGACTACACTCTCTTTTCTCTCAA 156049
QY 1922 GCGGAGCGCAAGGCCCTGAACTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 1981
Db 156048 GCGGAGCGCAAGGCCCTGAACTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 155989
QY 1982 TCCG-CCTTAGTCCCGCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 2040
Db 155988 TCCGCGCTTAGTCCCGCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 155929
QY 2041 CTCACCTCTACCCAAACCGCTAGGTTCCCGGTCGCCAGTACAGTCTGTATCAAAACCCACGA 2100
Db 155928 CTCACCTCTACCCAAACCGCTAGGTTCCCGGTCGCCAGTACAGTCTGTATCAAAACCCACGA 155869
QY 2101 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTTAGAAATATAGGTCTCTCTCC 2160
Db 155868 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTTAGAAATATAGGTCTCTCTCC 155809
QY 2161 CAGAATCCAGCGCGCCCAATGGAACCTCACGCTGGGTCTTAATTACCACTCTTTAAAGG 2220
Db 155808 CAGAATCCAGCGCGCCCAATGGAACCTCACGCTGGGTCTTAATTACCACTCTTTAAAGG 155749
QY 2221 CCCAGCCCTAGAAACCCAGCTCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 155748 CCCAGCCCTAGAAACCCAGCTCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 155689
QY 2281 TACTCAGGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCAGGGAGCAC 2340
Db 155688 TACTCAGGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCAA-GGAGCAC 155630
QY 2341 AGGTCCCACTCCAGCCCGGGAGCCTAGGACCACTCAGCCCTAGGAGTATATTTCCGC 2400
Db 155629 AGGTCCCACTCCAGCCC-GGAGGCTAGGACCACTCAGCCCTAGGAGTATATTTCCGC 155571
QY 2401 ACTTCAGAATTCATATCTTGGAAATCCAAGCTCCCTGCCCAATAAATTAAGTTCAGTCTCTGC 2460
Db 155570 ACTTCAGAATTCATATCTTGGAAATCCAAGCTCCCTGCCCAATAAATTAAGTTCAGTCTCTGC 155511
QY 2461 TTCCAGAATTTGGAAATCCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAAAC 2520
Db 155510 -TCCAGAATTTGGAAATCCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAAAC 155452
QY 2521 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGAGCAACCTGGCCCGG 2580
Db 155451 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGAGCAACCTGGCCCGG 155392
QY 2581 ATCAGACGAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCAGGCCCATCTCG 2640
Db 155391 ATCAGACGAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCAGGCCCATCTCG 155332
QY 2641 GAGCATCCCGGAGGAATCTGCAGAGGGGTTAGGAGTGGGTGACAGAGCCTGATCTCTT 2700
Db 155331 GAGCATCCCGGAGGAATCTGCAGA-GGGTTAGGAGTGGGTGACAGAGCCTGATCTCTT 155273

QY 2701 CCTGTTTCTACATAGATTTATTTTTCAGTTCACAAGAAAGATCAATACATTTTC11AAAA 2760
Db 155272 CCTGTTTCTACATAGATTTATTTTTCAGTTCACAAGAAAGATCAATACATTTTGT1AAAA 155213
QY 2761 AAAAAAAA 2770
Db 155212 AAAATATAA 155203

Search completed: October 14, 2001, 01:32:59
Job time: 20395 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 12, 2001, 18:18:32 ; Search time 27.54 Seconds  
           (without alignments)  
          2930.504 Million cell upd
```

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLPLRSIDVV.....KVSTPSDVTLSFLKRRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Database : SPTREMBL_16:*

```
1: sp_archea:*
```

```
2: sp_bacteria:*
```

```
3: sp_fungi:*
```

4: sp_human: *

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc: *

8: sp_organelle:

g:: sp_phage: *

```

10: sp plant:*
sp=prage:

```

```
10: sp_prauc:
11: sp_rocnet: *
```

```

11: sp_routenc:
12: sp_unclassified:*

```

```

12: sp_unclassified.*
13: sn vertebrate.*

```

```
13: sp_vertebrat
14: sp_virus:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	o				
1	3083	96.8	610	11	Q9WVR8		Q9WVR8 rattus norv
2	2109.5	66.2	617	13	Q9IAA9		Q9IAA9 brachydanio
3	2095.5	65.8	617	13	Q9PU64		Q9PU64 brachydanio
4	1800	56.5	341	4	Q9UE24		Q9UE24 homo sapien
5	1153.5	36.2	751	5	Q9NRX0		Q9NRX0 drosophila
6	1124.5	35.3	813	5	Q9VM47		Q9VM47 drosophila
7	734	23.0	4	146	4	Q9GZ05	Q9GZ05 homo sapien
8	139	4.4	531	5	Q24426		Q24426 drosophila
9	129	4.1	566	5	Q9VF97		Q9VF97 drosophila
10	128.5	4.0	609	13	Q13027		Q13027 xenopus lae
11	126.5	4.0	813	4	Q75229		Q75229 homo sapien
12	126	4.0	397	5	Q9GRQ7		Q9GRQ7 leishmania
13	126	4.0	545	11	O55206		O55206 mus musculus
14	125.5	3.9	4880	11	Q9JLT1		Q9JLT1 rattus norv
15	125.5	3.9	5085	11	Q9JKS6		Q9JKS6 rattus norv
16	124.5	3.9	992	14	Q69119		Q69119 human herpe
17	124	3.9	358	4	Q08805		Q08805 homo sapien
18	124	3.9	382	4	O00599		O00599 homo sapien
19	124	3.9	1320	2	Q9RJY6		Q9RJY6 streptomyce

ALIGNMENTS

```

RESULT 1
Q9WVR8
ID Q9WVR8 PRELIMINARY; PRT; 610 AA.
AC Q9WVR8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE MENIN.
GN MEN1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=WISTAR; TISSUE=BRAIN;
RA Maruyama K., Tsukada T., Hosono T., Ohkura N., Kishi M., Honda M.,
RA Nara-Ashizawa N., Nagasaki K., Yamaguchi K.;
RT "Molecular cloning and tissue distribution of mRNA for rat menin.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=99453728; PubMed=10524203;
RA Karges W., Maier S., Wisemann A., Dralle H., Dosch H.M., Boehm B.O.;
RT "Primary structure, gene expression and chromosomal mapping of rodent
RT homologs of the MEN1 tumor suppressor gene.";
RL Biochim. Biophys. Acta 1446:286-294(1999).
DR EMBL; AB023400; BA082134.1; -
DR EMBL; AF130370; AAF01354.1; -
DR EMBL; AF130369; AAF01353.1; -
SQ SEQUENCE 610 AA; 67335 MW; F9D2078B02363E32 CRC64;

Query Match 96.8%; Score 3083; DB 11; Length 610;
Best Local Similarity 97.0%; Pred. No. le=234;
Matches 592; Conservative 5; Mismatches 13; Indels 0; Gaps

Qy 1 MGLKAAQKTLPLRSIDDVWVRLFAALGREGPEPDVLLSLVLGVFVEHFLAVNRVIPTNVPE 60
Db 1 MGLKAAQKTLPLRSIDDVWVRLFAALGREGPEPDVLLSLVLGVFVEHFLAVNRVIPTNVPE 60
Qv 61 LTFQSPADPPGGTLTYFPVADLSITIAALYARFTAOIRGAVDLSLYPREGGVSSRELVRK 120

```



```
Db 1 MGIRSTQKKHFLRGIDGVQVLFSELSNPEPDLALLSLVLGFEVHFLAVNRVVPVNPVG 60
QY 61 LTFQSPAPDPGGLTYFPVADLSITAAIYARETAQIRGAVDLSLYPREGGVSSRELKVK 120
Db 61 VRFEP- --OPSLNCFPTVELNLSALYERMAQIRGAVDMSQTKRPSGASSRELKVK 117
QY 121 VSDIWNLSRSYFKDRAHIQSLFSPITCKLDSSGAVAFVVGACQALGLRDVHLALSED 180
Db 118 VSDIWNLSRSYFKDRAHIQSLFSLITCTKLDSSGAVAFVAAACQVGLKDVHLALSED 177
QY 181 HAWVPGPNGEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSMCRDKMEVAFVW 240
Db 178 HAWVPGKGEETAETVTHGKGNEDRRGQTVSAGINERSWLYLKGSMKCNRMMEVAFVW 237
QY 241 CAINPSIDLHTSLELLOLQOQLLWLLYDLGLHLYRPMALGNLADLEELEPTGPRDPLT 300
Db 238 CAINPSIDLHTSTELIQOQLLWLLYDRGLERYPMAMGTALADEDEPTGKESPLS 297
QY 301 LYHKGASAKTYRDBHIYPYMYLAGYHCRNVRNREALQAWADATVIOQNYNCREDEEI 360
Db 298 IHLKAVESAKYNNENHIYPYMYLAGYHYHREVREALGSWAEAAASVMQDNYNCREDEEI 357
QY 361 YKEFFPVANDVTPNLLKEAASLLEA--GEERPEQSGTOSQ--SALQDPECFAHLRLFYD 418
Db 358 YKEFFDIANDGPTLLKETAASAESEGTDEAEKEDQPREAALTAALQDPECFAHLRLFYD 417
QY 419 GICKWEESGPTVPLHVGWATFVLOSIGRFEQGVQKRVIVSREAEAAAEPEWGEAREG 478
Db 418 GICKWEESGPTVPLHVGWATYLVLOSIRFDAQIRQKVSITTKDTEPVDODDSSDREG 477
QY 479 RRRGPRRESKPEEP-----PPPKPA-----LDKGLGTGGAGVSGPRPKPGTGTAGTARG 528
Db 478 RRRGPRRESKLEDPAGSASPALPAONQNPVKKVG--GEG---GRRSSASTRGREADG 533
QY 529 P-EGGSTAQVPAPAAAPPE--GPVLTQSEKMKMKELLVANKINSATKLQTLTAOSQVQ 586
Db 534 KNEPSPSPISQOPAVQGVGVVFFHSEKMKMKELLSAKINSATKLQTLTAOSQVQ 593
QY 587 MKQKVSTPSTPSTLTLFLKQRKGL 610
Db 594 MKRQKPTASGDYTLSPMKRPRKTL 617

RESULT 4
Q9UE24 PRELIMINARY; PRT; 341 AA.
AC Q9UE24:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE MRNA 3'-REGION (UNKNOWN FUNCTION) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001089; PubMed=9341881;
RA Kedra D., Seroussi E., Fransson I., Trifunovic J., Clark M.,
RA Lagercrantz J., Blennow E., Mehlin H., Dumanski J.;
RT "The germinal centre kinase gene and a novel CXC25-like gene are
RL located in the vicinity of the PYGM gene on 11q13."
RL Hum. Genet. 100:611-619(1997).
DR EMBL; Y12338; CAA73007.1; -.
FT NON_TER
SQ SEQUENCE 341 AA; 37723 MW; 01108F638C4CC64C CRC64;
```

Query Match 56.5%; Score 1800; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.8e-134;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 LGHLERYPMALGNLADLEELEPTGPRDPLTLTYHKGIASAKTYRDEHIYPVNYLAGYHC 329

```
Db 1 LGHLERYPMALGNLADLEELEPTGPRDPLTLTYHKGIASAKTYRDEHIYPVNYLAGYHC 60
QY 330 RRRNREALQAWADATVIOQNYNCREDEEIYKEFFEVDANDVTPNLLKKAASLLEAGEER 389
Db 61 RRRNREALQAWADATVIOQNYNCREDEEIYKEFFEVDANDVTPNLLKKAASLLEAGEER 120
QY 390 PGEQSGTOSQSALQDPECFAHLRLFYDGICKWEESGPTVPLHVGWATFVLOSIGRFEQ 449
Db 121 PGEQSGTOSQSALQDPECFAHLRLFYDGICKWEESGPTVPLHVGWATFVLOSIGRFEQ 180
QY 450 QVRQKRVIVSREAEAAAEPEWGEAREGRRGPRRESKPEEPKKPALDKGLGTGQG 509
Db 181 QVRQKRVIVSREAEAAAEPEWGEAREGRRGPRRESKPEEPKKPALDKGLGTGQG 240
QY 510 AVSGPRPKPGTGTAGTARGPEGGSTAQVPAPAAAPPEGVPVLTQSEKMKMKELLVATK 569
Db 241 AVSGPRPKPGTGTAGTARGPEGGSTAQVPAPAAAPPEGVPVLTQSEKMKMKELLVATK 300
QY 570 INSSAIKLQTLTAOSQVOMKKQKVSTPSTDYTLFLKQRKGL 610
Db 301 INSSAIKLQTLTAOSQVOMKKQKVSTPSTDYTLFLKQRKGL 341

RESULT 5
Q9NKX0 PRELIMINARY; PRT; 751 AA.
AC Q9NKX0:
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE MENIN.
GN MN1 OR MEN1 OR CG13778.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Maruyama K., Tsukada T., Honda M., Ohkura N., Nagasaki K., Kishi M.,
RA Nara-ashizawa N., Yamaguchi K.;
RT "Drosophila menin cDNA.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB040816; BAA94341.1; -.
DR FlyBase; FBgn0031885; Men1.
SQ SEQUENCE 751 AA; 81674 MW; D18EA61238C14115 CRC64;
```

Query Match 36.2%; Score 1153.5; DB 5; Length 751;
Best Local Similarity 38.1%; Pred. No. 2.1e-82;
Matches 272; Conservative 91; Mismatches 195; Indels 155; Gaps 15;

```
QY 9 TLFPLRSIDVVRLFAEL--GREEDLVLSLVGVFVHFLAVNRVITPNVPELTQPSQ 67
Db 10 SLFPLKSTADVINLFRALTSIGTEPDLTLLSIVGVVIELSTTGEAAQ-----AAQAAA 64
QY 68 APDPGGLTY-----FPVADLSITAAIYARETAQIRGAVDLSLYPREGGV 112
Db 65 AAVAAGDISOATGGNDIINGNSVPPFVTHLIGLYKKFOT--ILSVBE---KPKPHRQ 120
QY 113 SSRELKVKYSDVWNSLSRSYFKDRAHIQSLFSGITGTCKLDSSGAVAFVVGACQALGLRD 172
Db 121 ATREVIKKYSDVWNSLSRSYFKDRAHLQNLVSYLSGNKLDCCFVALAVVAGCOLLYK 180
QY 173 VHLALSEDHAWVYFGPNGEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSMYRCDR 232
Db 181 VRLAISEDHAWVYFGQKRVETIEVTHGKSGEDKRGDIPRGIESGWSLYLGLGAVVCR 240
QY 233 KMEYAVWCAINPSIDLHTSLELLOLQOQLLWLLYDLGLHLYRPMALGNLADLEELEPT 292
Db 241 GMEVAAICAAALNISLTNSDCVEAEVLEQQQLLWLLYDLGLHLYRPMALGNLADLEELEPT 300
```

```
QY 293 PGRDPDLTYLHKGIASAKTYRDEHIYPMYLAGYHCRNRNVRALQAMADTATVIQDN 352
Db 301 HPSTSCBOLYREATESARTYRHHVHPYTYQNGYNNRLKYRDAFAAANAADVIRLYT 360
QY 353 Y-CREDEEIKYKEFEVANDVIPNLLKEAASLLEAGEERPEQSGTOSQG-----SALODP 407
Db 361 YQCRDDEEIKYKELLDIANELIPYMK-----TESSGHSAIRLDS 401
QY 408 ECFAPHLRFDYGICKWPEGSPPTVLHVGMATFLVQSLGRPEGOVRKVR-----457
Db 402 EVFANLLRFYDYGICQWEDSLTPILHIGWAKPLVNNITKFDYDIRSQVVKLPEDLEAQ 461
QY 458 -----VSREA-EAAAEAPWEGEAREGRRR-----GPRRESKPEPP-----493
Db 462 AKAEARAEQEAKEAKESKAAGSEAMEGNN--RMAKKEKSKNSLPTTLADLTAACG 519
QY 505 -----GTGQ-----GAVSGPP-----RKPPCTV 522
Db 520 EKILNPDLQGGQPFADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 579
QY 523 AGTARGPEGSTAQVPAPASPPPEG-----548
Db 580 TTNATTTSNGSGTSVQLPVSEANNAQAOQVQINDOLGKPOHKEAKKETSDDYDPFE 639
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSOVMKOKVSTPS 596
Db 640 IMLKRPVITLYSOKMKGLKLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 692

RESULT 6
QYVM47 PRELIMINARY; PRT; 813 AA.
AC QYVM47;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
DE CG13778 PROTEIN.
GN MN1 OR CG13778.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Prterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Boutler H., Brokstein P., Brottier P.,
RA Chert J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista A.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hardek N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
```

```
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupers M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun Y.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003616; AAF52480.1; -.
DR FlyBase; FBgn0031885; Mn1.
DR InterPro; IPR001609; -.
DR ProDom; PD000355; -.
SQ SEQUENCE 813 AA; 88384 MW; 4B7ED54B4BLC2AAF CRC64;

Query Match 35.3%; Score 1124.5; DB 5; Length 813;
Best Local Similarity 35.4%; Pred. No. 4.5e-80;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

QY 9 TLPLRSIDDDVRLFAEL-GREEPDVLVLSVLGFVEHFLAVNRVPTNVPELTQPSP 67
Db 10 SLFPLKSTADVINFRRALTSGIEPDLTLLSYVGVYELSLTTGEAAQ-----AAQAAA 64
QY 68 APDPGGLTY-----FPVADLSIIAALYARFTAQIRGAVDLSLYPREGGV 112
Db 65 AAVAAGDISAQTGGNDIIMGNSVPPVPTVTHELAGYKKFQT-ILSVBE--KKPKPHRQ 120
QY 113 SSRELKVKSDVIWNSLSRYFKDRAHQISLFSFITGKLDSSGVAFVAVGACQALGRD 172
Db 121 ATREVKKVSDVIWNSLSRYSYKDRALHQLNLYSLGKLDCEGVALVAVAGCOLLYGCR 180
QY 173 VHLALEDHAWYVFGNGEGTAEVTHHGKGNEDRRGOTVAGVAERSWLKSGSYMCDR 232
Db 181 VRLAISEDHAWYVFGKRVETIEVTHGKGSDEKRGDIRPGEISGWLVLGGLVAVVCR 240
QY 233 KMEVAFMVCAINPSIDLHTLSLELQLOQLLWLLYDLGLHLYRYPMALGNLDELEPT 292
Db 241 GMEVAICAALNLSLTSNSDCVEAELOQQLLWLLYDLGLHLYRYPMALGTLGELEIHT 300
QY 293 PGRDPDLTYLHKGIASAKTYRDEHIYPMYLAGYHCRNRNVRALQAMADTATVIQDN 352
Db 301 HPSTSCBOLYREATESARTYRHHVHPYTYQNGYNNRLKYRDAFAAANAADVIRLYT 360
QY 353 Y-CREDEEIKYKEFEVANDVIPNLLKEAASLLEAGEERPEQSGTOSQG-----SALODP 407
Db 361 YQCRDDEEIKYKELLDIANELIPYMK-----TESSGHSAIRLDS 401
QY 408 ECFAPHLRFDYGICKWPEGSPPTVLHVGMATFLVQSLGRPEGOVRKVR-----457
Db 402 EVFANLLRFYDYGICQWEDSLTPILHIGWAKPLVNNITKFDYDIRSQVVKLPEDLEAQ 461
QY 458 -----VSREA-EAAAEAPWEGEAREGRRR-----GPRRESKPEPP-----493
Db 462 AKAEARAEQEAKEAKESKAAGSEAMEGNN--RMAKKEKSKNSLPTTLADLTAACG 519
QY 494 -----GTGQ-----GAVSGPP-----RKPPCTV 522
Db 520 EKILNPDLQGGQPFADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 579
QY 523 AGTARGPEGSTAQVPAPASPPPEG-----548
Db 642 TTNATTTSNGSGTSVQLPVSEANNAQAOQVQINDOLGKPOHKEAKKETSDDYDPFE 701
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSOVMKOKVSTPS 596
```

Db 702 IMKRPVITLYSOKMKGLKDLALLAEKLNTHAISLQVTAQSVASRKVRGVKHS 754

```
RESULT 7
Q9CZQ5
ID Q9GZQ5 PRELIMINARY; PRT; 146 AA.
AC Q9GZQ5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MENIN (FRAGMENT).
GN MENI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20484185; PubMed=11027505;
RA Khodaei-O'Brien S., Zabłowska B., Fromaget M., Bylund L., Weber G.,
RA Gaudray P.;
RL Biochem. Biophys. Res. Commun. 276:508-514(2000).
DR EMBL; AJ297489; CAC14133.1; -.
DR EMBL; AJ297485; CAC14129.1; -.
DR EMBL; AJ297486; CAC14130.1; -.
DR EMBL; AJ297487; CAC14131.1; -.
DR EMBL; AJ297488; CAC14132.1; -.
FT NON_TER 146
SQ SEQUENCE 146 AA; 16145 MW; BD7EDFF9A7AE03C7 CRC64;
```

```
Query Match 23.0%; Score 734; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 2.3e-50;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVFVHFLAVNRVPTNPVE 60
Db 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVFVHFLAVNRVPTNPVE 60

QY 61 LTFQSPAPDPGGTLYFPVADLSIIAALYARFTAQIRGAVDSLYPREGGVSSRELVRK 120
Db 61 LTFQSPAPDPGGTLYFPVADLSIIAALYARFTAQIRGAVDSLYPREGGVSSRELVRK 120

QY 121 VSDVINWNSLSRSYFKDRAHQSLFSF 146
Db 121 VSDVINWNSLSRSYFKDRAHQSLFSF 146
```

```
RESULT 8
Q24426
ID Q24426 PRELIMINARY; PRT; 531 AA.
AC Q24426;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE TROPOMYOSIN ISOFORM 33.
GN TM1 OR CG4898.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84205681; PubMed=6202423;
RA Karlik C.C., Mahaffey J.W., Coutu M.D., Fyrberg E.A.;
RT "Organization of contractile protein genes within the 88F subdivision
RT of the D. melanogaster third chromosome.";
RL Cell 37:469-481(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064486; PubMed=3097506;
```

```
RA Karlik C.C., Fyrberg E.A.;
RT "Two Drosophila melanogaster tropomyosin genes: structural and
RT functional aspects.";
RL Mol. Cell. Biol. 6:1965-1973(1986).
DR EMBL; K02621; AAA28968.1; -.
DR EMBL; M12840; AAA28968.1; JOINED.
DR EMBL; L00355; AAA28968.1; JOINED.
DR EMBL; L00356; AAA28968.1; JOINED.
DR EMBL; L00357; AAA28968.1; JOINED.
DR EMBL; L00358; AAA28968.1; JOINED.
DR EMBL; L00359; AAA28968.1; JOINED.
DR EMBL; L00360; AAA28968.1; JOINED.
DR EMBL; L00362; AAA28968.1; JOINED.
DR HSSP; P04002; 1ATF.
DR FlyBase; FBgn003721; Tm1.
DR InterPro; IPR000104; -.
DR InterPro; IPR000533; -.
DR Pfam; PF00261; Tropomyosin; 2.
DR PRINTS; PR00308; ANTIFREEZEI.
DR PRINTS; PR01217; PRICHEXTENSN.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PROSITE; PS00326; TROPOMYOSIN; 1.
KW Alternative splicing.
SQ SEQUENCE 531 AA; 56047 MW; 2618A715E20EE4DF CRC64;
```

```
Query Match 4.4%; Score 139; DB 5; Length 531;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 105; Conservative 50; Mismatches 194; Indels 144; Gaps 20;

QY 114 SRELKVKSDVINWNSLSRSYFKDRAHQSLFSFITGKTLDSSGVAF-----AVVGACQAL 168
Db 43 ARQLQKKI-QTVENELDQI-----QEAALTIVTG-KLEEKNKALQNAESEAALNRRI 92

QY 169 GLRDVHLALSEDHAWVFGPNGBQTAETVHWGKGNDRGQTVNAGVARSWLYLKGSYM 228
Db 93 QLLEEDLERSEERSASAI---QLAAEASQSADESERARKILENRALADEERMDALENQL 148

QY 229 RCDRKNVAFWCAINPSIDLHTDLSLELLQLOOKLLWLLYDLGLHLYPMALGNLADLEE 288
Db 149 K-----EARFLAEEADKKYD-----EVARKLAWVEADLERAEERAMV-----EADLER 191

QY 289 LEPTGCRPDPLTLYHKGIASAKTYRDEHYIPYMYLAGYHCRNRNVREALQANADTATVI 348
Db 192 AEERAEQ-----GENKIVELEEE-----LRLVGNLKSLEYSE----- 224

QY 349 QDYNVCREDEEYKEFFEVDANDVIPNLKKEAASLLLEAGEERPGEOGQTSQGSALQD-- 406
Db 225 ---EKANQREEEYKNQIKTLN----TRLKEA-----EAAEFAERSVOKLQKEVDRLEDDL 273

QY 407 -----PECFALLRFDYGICKWEBS---PTPVLHVGNWATFLVOSLSGRFEG 449
Db 274 IVEKERYCMIGDSLDEAFVDLKLGLPEFF--WNPKNPKPTPKL-----PTPTPEELAAAME 337

QY 450 QVRQKVRIVSRGAEAAAEAEPEWGEAREGRRRGRPRRESKP-EP PPPPKKPKALDKGLGTGQ 508
Db 328 ARAAAEAAAAAEAAAEAAAEAGACADGAPAPGEGAKAPEPTPKPEP----- 377

QY 509 GAVSGPPRPKPPGTVAGTARGEGGTAQV----- 537
Db 378 ---TPPPPPPPPEYSIDLPPGEAGVYKVNVEPPPPGSEPEVPAAGEGAAPAEAGAAP 434

QY 538 PAPAASPPPEGPV 550
Db 435 PAEGAAPPAEGAV 447

RESULT 9
Q9VF97
ID Q9VF97 PRELIMINARY; PRT; 566 AA.
AC Q9VF97;
DT 01-MAY-2000 (Tremblrel. 13, Created)
```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE TMI PROTEIN.
 GN TMI OR CG4898.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003708; AAF55164.1; -
 DR HSP; P04002; 1ATF.
 DR FlyBase; FBgn003721; Tml.
 DR InterPro; IPR00104; -
 DR InterPro; IPR000533; -
 DR InterPro; IPR002965; -
 DR Pfam; PF00261; Tropomyosin; 2.
 DR PRINTS; PR00308; ANTIFREEZE1.
 DR PRINTS; PR01217; PRICHEXTENS.
 DR PRINTS; PR00194; TROPOMYOSIN.
 DR PROSITE; PS00326; TROPOMYOSIN; 1.
 SQ SEQUENCE 566 AA; 60449 MW; 2AF07CFB0DA2032E CRC64;

Query Match 4.18; Score 129; DB 5; Length 566;
 Best Local Similarity 25.4%; Pred. No. 0.086;
 Matches 62; Conservative 14; Mismatches 90; Indels 78; Gaps 9;

QY 358 EEIYKEFEVANDVIPNLLKEAASLLGAGEERPGESQSGTOSGALQD----- 406
 DB 266 EEEYKNQIKTLN----FRLKEA----EAAEEAERSVQKQKQVDRLEDDLLIVEKRYCM 317

QY 407 -----PECFALLRFRYDGIKWEES---PTVLHVGMATFLVQSLGRFCQVQRKVRIV 458
 DB 318 IGDLSDEAFVDLIKGLPEF--WNRNPKPPTPKL-----PTPTPEELAAEMEAARAAEAAA 371
 QY 459 SRRAEAAEAEPEWGEAREGRRRGPRRESKP-EEPPPPKPPALDKGLCTGGAVSGPPRK 517
 DB 372 AAEAAEAAAAAGEAGPDGAPAAPEGEKAPAKEPTPKPE-----TPPPPP 418
 QY 518 PPGTVAGTARGPGGSTAQV-----PAPAAASPPP 546
 DB 419 PPPEYSIDLPPGEAEVYVKNYEPPEPPGSEPEVPAAEGEAPAAECAAAPPAEAPPA 478
 QY 547 EGPV 550
 DB 479 EGAV 482
 RESULT 10
 ID 013027 PRELIMINARY; PRT; 609 AA.
 AC 013027;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE MITOTIC PHOSPHOPROTEIN 90.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97272720; PubMed=9115395;
 RA Stukenberg P.T., Lustig K.D., McGarry T.J., King R.W., Kuang J.,
 RA Kirschner M.W.;
 RT "Systematic identification of mitotic phosphoproteins."
 RL Curr. Biol. 7:338-348(1997).
 DR EMBL; U95102; AAC60123.1; -
 DR InterPro; IPR001026; -
 DR Pfam; PF01417; ENTH; 1.
 DR SMART; SM00373; ENTH; 1.
 SQ SEQUENCE 609 AA; 64589 MW; E65EC0CC954D2459 CRC64;
 Query Match 4.0%; Score 128.5; DB 13; Length 609;
 Best Local Similarity 23.7%; Pred. No. 0.1;
 Matches 71; Conservative 29; Mismatches 90; Indels 109; Gaps 14;
 QY 317 HIIPYMYLAGHCHNRNRVREALQAWADTAIV--IODYNYCREDEEIIYKEFEFVANDVIPN 374
 DB 64 HVYKAMTLMETLIKTGSERVAQCKENIYATQTLKDFYVDRO-----GKDGVN 113
 QY 375 LLKEAASLLLEAGEERPGESQSGTOSGALQDPCF----AHLRFYDGIKWEES--- 427
 DB 114 VREKAKOLV-----SLKDDERLKEERNAHLKTEK(AOT)STSSAS 155
 QY 428 ----PTVLHVGMATFLVQSLGRFGQVRKVRIVSRSEAAEAEPE----- 470
 DB 156 STLNPAPEGEQAWS-----QSSGEELQLALANSEKAEQVRAKPPVSEELQLAL 211
 QY 471 -----WGEEAR-----EGRRGPPRESKPEE-----PPPKPKP 498
 DB 212 SLSKEEHDKKEERIKRGDDLRLQMALESRRKGP---SKQEQQSSMLDLADVFSPPAPVAP 268
 QY 499 ALDKGLCTGGAVSGPPKPPGTVAGTARGPEGGSGTQVAPAPASPPPEG--PVLTFQS 555
 DB 269 TADP-----WGASAAPPADP---WAG-----GATPASVPANAAAPDPWCGPPVATGSS 313
 RESULT 11
 ID 075229 PRELIMINARY; PRT; 813 AA.
 AC 075229;

DR EMBL; AF227534; AAF63196.1; -
 DR InterPro; IPR000008; -
 DR InterPro; IPR001478; -
 DR InterPro; IPR002965; -
 DR Pfam; PF00168; C2; 2.
 DR PRINTS; PRL01217; PRICHEXTENS.
 DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
 DR PROSITE; PS00004; C2_DOMAIN_2; 2.
 DR SMART; SM00239; C2; 1.
 KW Matrix protein.
 SQ SEQUENCE 5085 AA; 552711 MW; 5A1BB543201A7450 CRC64;

Search completed: October 12, 2001, 18:21:53
Job time: 201 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:18:52 ; Search time 13.19 seconds
(without alignments)
1584.219 Million cell updates/sec

Title: US-09-380-337-2

Perfect score: 3185

Sequence: 1 MGLKAAQKTLFLRLSIDDV.....KVSTPSDYTLFLKRRKGL 610

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3172.5	99.6	615	1 MEN1_HUMAN	O00255 homo sapien
2	3049.5	95.7	611	1 MEN1_MOUSE	O88359 mus musculus
3	141.5	4.4	518	1 TPM4_DROME	P49455 drosophila
4	132	4.1	1516	1 CALH_HUMAN	P39060 homo sapien
5	125.5	3.9	1025	1 CAL6_MOUSE	O04857 mus musculus
6	124.5	3.9	992	1 EBN6_EBV	P03204 epstein-bar
7	123	3.9	331	1 PRP1_HUMAN	P04280 homo sapien
8	122.5	3.8	1790	1 SEPA_EME1	P78621 emericella
9	119	3.7	172	1 PRP2_RAT	P10164 rattus norv
10	118	3.7	705	1 PRP2_HUMAN	P17600 homo sapien
11	117.5	3.7	251	1 PRP2_HUMAN	P02812 homo sapien
12	117.5	3.7	1581	1 PRPB_HUMAN	O15648 h peroxisom
13	117	3.7	308	1 MACS_MOUSE	P26645 mus musculus
14	116	3.6	704	1 SYNL_RAT	P09951 rattus norv
15	116	3.6	726	1 ADDB_HUMAN	P35612 homo sapien
16	116	3.6	860	1 AREA_PENRO	O13508 penicillium
17	116	3.6	2774	1 MAPA_RAT	P34926 rattus norv
18	115.5	3.6	247	1 PRP4_HUMAN	P10163 homo sapien
19	114	3.6	492	1 MEC2_RAT	O00566 rattus norv
20	113.5	3.6	479	1 ACMA_HUMAN	P08173 homo sapien
21	113.5	3.6	1838	1 CAL5_HUMAN	P20908 homo sapien
22	113	3.5	484	1 MEC2_MOUSE	O922d6 mus musculus
23	113	3.5	706	1 SYNL_BOVIN	P17599 bos taurus
24	112.5	3.5	384	1 AIPL1_HUMAN	O9n2n9 homo sapien
25	111.5	3.5	861	1 DYN1_MOUSE	P39053 mus musculus
26	111.5	3.5	2453	1 NCRI_MOUSE	O60974 mus musculus
27	111	3.5	276	1 PRPL_HUMAN	P10162 homo sapien
28	111	3.5	520	1 WASP_MOUSE	P70315 mus musculus
29	111	3.5	1888	1 CAL1_CHICK	P32018 gallus gall
30	110.5	3.5	549	1 AGLA_RHIME	O923i8 rhizobium m
31	110	3.5	174	1 PRPP_HUMAN	P81489 homo sapien
32	110	3.5	479	1 ACMA_MOUSE	P32211 mus musculus
33	110	3.5	801	1 BRD2_HUMAN	P25440 homo sapien

RESULT	ID	MEN1_HUMAN	STANDARD;	PRT;	615 AA.
34	110	3.5	857	1 NFM_CHICK	P16053 gallus gall
35	110	3.5	2220	1 YICI_HUMAN	O9y610 homo sapien
36	109.5	3.4	772	1 YICI_ECOLI	P31434 escherichia
37	109	3.4	670	1 SYNI_MOUSE	O88935 mus musculus
38	109	3.4	783	1 FYB_HUMAN	O15117 homo sapien
39	108.5	3.4	864	1 DYN1_HUMAN	O05193 homo sapien
40	108.5	3.4	865	1 NRFA_PENUR	O92269 penicillium
41	108	3.4	342	1 ME18_MOUSE	P23798 mus musculus
42	108	3.4	689	1 L100_ADECC	O65957 canine aden
43	108	3.4	2205	1 POLN_RUBVT	P13889 rubella vir
44	107.5	3.4	384	1 VASP_CANFA	P50551 canis famil
45	106.5	3.3	308	1 MACS_RAT	P30009 rattus norv

ALIGNMENTS

RESULT	ID	MEN1_HUMAN	STANDARD;	PRT;	615 AA.
34	110	3.5	857	1 NFM_CHICK	P16053 gallus gall
35	110	3.5	2220	1 YICI_HUMAN	O9y610 homo sapien
36	109.5	3.4	772	1 YICI_ECOLI	P31434 escherichia
37	109	3.4	670	1 SYNI_MOUSE	O88935 mus musculus
38	109	3.4	783	1 FYB_HUMAN	O15117 homo sapien
39	108.5	3.4	864	1 DYN1_HUMAN	O05193 homo sapien
40	108.5	3.4	865	1 NRFA_PENUR	O92269 penicillium
41	108	3.4	342	1 ME18_MOUSE	P23798 mus musculus
42	108	3.4	689	1 L100_ADECC	O65957 canine aden
43	108	3.4	2205	1 POLN_RUBVT	P13889 rubella vir
44	107.5	3.4	384	1 VASP_CANFA	P50551 canis famil
45	106.5	3.3	308	1 MACS_RAT	P30009 rattus norv

RP VARIANTS FMEN1.
 RX MEDLINE=98349969; PubMed=9683585;
 RA Giraud S., Zhang C.X., Serova-Sinilnikova O., Wautot V., Salandre J.,
 RA Buisson N., Waterlot C., Bauders C., Porchet N., Aubert J.-P., Emy P.,
 RA Cadot G., Delamer B., Chabre O., Niccoli P., Lepat F., Duron F.,
 RA Emperauger B., Coudard P., Goudet P., Sarfati E., Riou J.-P.,
 RA Guichard S., Rodier M., Meyrier A., Caron P., Vantghem M.-C.,
 RA Assayag M., Peix J.-L., Pugeat M., Rohmer V., Vallotton M., Lenoir G.,
 RA Gaudray D., Proye C., Conte-Devolx B., Chanson P., Shugart Y.Y.,
 RA Goldgar D., Murat A., Calender A.;
 RT "Germline mutation analysis in patients with multiple endocrine
 RT neoplasia type 1 and related disorders.";
 RL Am. J. Hum. Genet. 63:455-467(1998).
 RP [6]
 RP VARIANT FIHP LYS-260.
 RX MEDLINE=99011276; PubMed=9792884;
 RA Teh B.T., Esapa C.T., Houlston R., Grandell U., Farnebo F.,
 RA Nordenskjold M., Pearce C.J., Carmichael D., Larsson C., Harris P.E.;
 RT "A family with isolated hyperparathyroidism segregating a missense
 RT MEN1 mutation and showing loss of the wild-type alleles in the
 RT parathyroid tumors.";
 RL Am. J. Hum. Genet. 63:1544-1549(1998).
 RP [7]
 RP VARIANT FIHP GLU-189.
 RX MEDLINE=99057176; PubMed=9843042;
 RA Fujimori M., Shirahama S., Sakurai A., Hashizume K., Hama Y., Ito K.,
 RA Shingu K., Kobayashi S., Anano J., Fukushima Y.;
 RT "Novel V184E MEN1 germline mutation in a Japanese kindred with
 RT familial hyperparathyroidism";
 RL Am. J. Med. Genet. 80:221-222(1998).
 RP [8]
 RP VARIANTS FMEN1.
 RX MEDLINE=98334342; PubMed=9671267;
 RA Agarwal S.K., Debelenko L.V., Kester M.B., Guru S.C., Manickam P.,
 RA Olufemi S.-E., Skarulis M.C., Heppner C., Crabtree J.S.;
 RA Lubensky I.A., Zhuang Z., Kim Y.S., Chandrasekharappa S.C.,
 RA Collins F.S., Liotta L.A., Spiegel A.M., Burns A.L., Emmert-Buck M.R.,
 RA Marx S.J.;
 RT "Analysis of recurrent germline mutations in the MEN1 gene encountered
 RT in apparently unrelated families.";
 RL Hum. Mutat. 12:75-82(1998).
 RP [9]
 RP VARIANT FMEN1 ILE-135 AND LYS-364.
 RX MEDLINE=98410971; PubMed=9740255;
 RA Boeni R., Vormeyer A.O., Pack S., Park W.-S., Burg G., Hofbauer G.,
 RA Darling T., Liotta L., Zhuang Z.;
 RT "Somatic mutations of the MEN1 tumor suppressor gene detected in
 RT sporadic angiofibromas";
 RL J. Invest. Dermatol. 111:539-540(1998).
 RP [10]
 RP VARIANTS FMEN1 LYS-119 DEL AND GLN-171--LEU-173 DEL.
 RX MEDLINE=98419173; PubMed=9747036;
 RA Sakurai A., Shirahama S., Fujimori M., Katai M., Itakura Y.,
 RA Kobayashi S., Anano J., Fukushima Y., Hashizume K.;
 RT "Novel MEN1 gene mutations in familial multiple endocrine neoplasia
 RT type 1.";
 RL J. Hum. Genet. 43:199-201(1998).
 RP [11]
 RP VARIANT FMEN1 GLY-45.
 RX MEDLINE=99048878; PubMed=9832038;
 RA Sato M., Matsubara S., Miyauchi A., Ohye H., Imachi H., Murao K.,
 RA Takahara J.;
 RT "Identification of five novel germline mutations of the MEN1 gene in
 RT Japanese multiple endocrine neoplasia type 1 (MEN1) families.";
 RL J. Med. Genet. 35:915-919(1998).
 RP [12]
 RP VARIANTS FMEN1 TRP-39; TYR-177; ASP-184; PRO-269 AND PRO-272.
 RX MEDLINE=99103464; PubMed=9886389;
 RA Poncin J., Abs R., Velkeniers B., Bonduelle M., Abranowicz M.,
 RA Legros J.-J., Verloes A., Meurisse M., van Gaal L., Verellen C.,
 RA Koulischer L., Beckers A.;
 RT "Mutation analysis of the MEN1 gene in Belgian patients with multiple
 RT endocrine neoplasia type 1 and related diseases";

RL Hum. Mutat. 13:54-60(1999).
 RP [13]
 RP VARIANTS MEN1 ASP-161 AND ARG-246.
 RX MEDLINE=99188881; PubMed=10090472;
 RA Mutch M.G., Dilley W.G., Sanjurjo F., Debenedetti M.K., Doherty G.M.,
 RA Wells S.A. Jr., Goodfellow P.J., Laimore T.C.;
 RT "Germline mutations in the multiple endocrine neoplasia type 1 gene:
 RT evidence for frequent splicing defects.";
 RL Hum. Mutat. 13:175-185(1999).
 CC -1- FUNCTION: NOT KNOWN.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -1- DISEASE: DEFECTS IN MEN1 IS THE CAUSE OF FAMILIAL MULTIPLE
 CC ENDOCRINE NEOPLASIA TYPE 1 (FMEN1); WERNER SYNDROME, AN AUTOSOMAL
 CC DOMINANT DISORDER CHARACTERIZED BY TUMORS OF THE PARATHYROID
 CC GLANDS, GASTRO-INTESTINAL ENDOCRINE TISSUE, THE ANTERIOR PITUITARY
 CC AND OTHER TISSUES. CUTANEOUS LESIONS AND NERVOUS-TISSUE TUMORS
 CC CAN EXIST. PROGNOSIS IN FMEN1 PATIENTS IS RELATED TO HORMONAL
 CC HYPERSECRETION BY TUMORS, SUCH AS HYPERGASTRINEMIA CAUSING SEVERE
 CC PEPTIC ULCER DISEASE (ZOLLINGER-ELLISON SYNDROME, ZES), PRIMARY
 CC HYPERPARATHYROIDISM, AND ACUTE FORMS OF HYPERINSULINEMIA.
 CC -1- DISEASE: DEFECTS IN MEN1 IS THE CAUSE OF FAMILIAL ISOLATED
 CC HYPERPARATHYROIDISM (FIHP OR HRPT1). FIHP IS AN AUTOSOMAL DOMINANT
 CC DISORDER CHARACTERIZED BY HYPERCALCEMIA, ELEVATED PARATHYROID
 CC HORMONE (PTH) LEVELS, AND UNIGLANDULAR OR MULTIGLANDULAR
 CC PARATHYROID TUMORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U93236; AAC51228.1; -;
 DR EMBL; U93237; AAC51229.1; -;
 DR EMBL; U93237; AAC51230.1; -;
 DR MIM; 131100; -;
 DR MIM; 145000; -;
 KW Nuclear protein; Disease mutation: Alternative splicing; Polymorphism.
 FT VARSPLIC 149 153 MISSING (IN SHORT ISOFORM).
 FT VARIANT 12 12 P-> L (IN FMEN1).
 FT VARIANT 22 22 /FTID=VAR_005425.
 FT L-> R (IN FMEN1).
 FT VARIANT 26 26 /FTID=VAR_005426.
 FT E-> K (IN PARATHYROID ADENOMA).
 FT VARIANT 39 39 /FTID=VAR_005427.
 FT L-> W (IN FMEN1).
 FT VARIANT 42 42 /FTID=VAR_005428.
 FT G-> D (IN FMEN1).
 FT VARIANT 45 45 /FTID=VAR_005429.
 FT E-> G (IN FMEN1).
 FT VARIANT 119 119 /FTID=VAR_005430.
 FT MISSING (IN FMEN1).
 FT VARIANT 135 135 /FTID=VAR_005431.
 FT K-> I (IN FMEN1).
 FT VARIANT 139 139 /FTID=VAR_005434.
 FT H-> D (IN FMEN1).
 FT VARIANT 139 139 /FTID=VAR_005432.
 FT H-> Y (IN FMEN1 AND SPORADIC MEN1).
 FT /FTID=VAR_005433.
 FT F-> V (IN FMEN1).
 FT VARIANT 144 144 /FTID=VAR_005436.
 FT G-> D (IN FMEN1).
 FT VARIANT 161 161 /FTID=VAR_008017.
 FT A-> P (IN FMEN1).
 FT VARIANT 165 165 /FTID=VAR_005437.
 FT A-> D (IN FMEN1).
 FT VARIANT 169 169 /FTID=VAR_005438.
 FT MISSING (IN FMEN1).
 FT VARIANT 171 173 /FTID=VAR_005439.
 FT R-> Q.

Query Match		99.6%;	Score 3172.5;	DB 1;	Length 615;
Best Local Similarity		99.2%;	Pred. No. 99-199;		
Matches 610;		Conservative	0;	Mismatches	0;
				Indels	5;
				Gaps	1;
QY	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVEHFHFLAVNRVPTNPVE	60		
Db	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVEHFHFLAVNRVPTNPVE	60		
QY	61	LTFQSPADPPGGLTYFFVADLSIIAALYARETAIRGAVDLSLYPREGVSSRELKVK	120		
Db	61	LTFQSPADPPGGLTYFFVADLSIIAALYARETAIRGAVDLSLYPREGVSSRELKVK	120		
QY	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITT-----GTLKDSGGVAFVAVGACQALGLRDVHL	175		
Db	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITGWSPTGTLKDSGGVAFVAVGACQALGLRDVHL	180		
QY	176	ALSEDHAWVFGNGBQTAEVTHWKGNEDEDRGQTVNAGVAERSWLYLKGSMRCRDKME	235		
Db	181	ALSEDHAWVFGNGBQTAEVTHWKGNEDEDRGQTVNAGVAERSWLYLKGSMRCRDKME	240		
QY	236	VAFMVCAINPSIDLHTDLSLELLOQLKLLWLLYDLGLHLERYPMALGNLADLLELEPTGR	295		
Db	241	VAFMVCAINPSIDLHTDLSLELLOQLKLLWLLYDLGLHLERYPMALGNLADLLELEPTGR	300		
QY	296	PDPLTYLHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCR	355		
Db	301	PDPLTYLHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCR	360		
QY	356	EBEIIYKEFEFVANDVIPNLLKEASLLBAGERPGEQSGQTSQSALQDPCEFAHLR	415		
Db	361	EBEIIYKEFEFVANDVIPNLLKEASLLBAGERPGEQSGQTSQSALQDPCEFAHLR	420		
QY	416	FYDGICKWEESGTPVLHVGMATFLVQSLGRREGQVRQKRVIRSREAAEAEPGEEA	475		
Db	421	FYDGICKWEESGTPVLHVGMATFLVQSLGRREGQVRQKRVIRSREAAEAEPGEEA	480		
QY	476	REGRRGRPRRESKPEPPPPKPKALDKGLTGQGVSGPPPKPPGTVAGTARGPEGSTA	535		
Db	481	REGRRGRPRRESKPEPPPPKPKALDKGLTGQGVSGPPPKPPGTVAGTARGPEGSTA	540		
QY	536	QVPAPASPPPEGPVLTFOSEKMKMKELLVATKINSIAIKLQLTQSQVQMKKQKVSPT	595		
Db	541	QVPAPASPPPEGPVLTFOSEKMKMKELLVATKINSIAIKLQLTQSQVQMKKQKVSPT	600		
QY	596	SDYTLFLKQRKGL 610			
Db	601	SDYTLFLKQRKGL 615			
RESULT 2					
MENI_MOUSE					
ID	MENI_MOUSE	STANDARD;	PRT;	611	AA.
AC	O88559;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	MENI.				
GN	MENI.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
SEQUENCE FROM N.A.					
RP	Medline=99039765; PubMed=9824159;				
RA	Stewart C., Parente F., Piehl F., Farnebo F., Quincey D., Silins G.,				
RA	Bergman L., Carle G.F., Lemmens I., Grimmond S., Xian C.Z., Khodai S.,				
RA	Teh B.T., Lagercrantz J., Siggers P., Calender A., van de Vem V.,				
RA	Kas K., Weber G., Hayward N., Gaudray P., Larsson C.				
RT	"Characterization of the mouse Menl gene and its expression during				
RT	development."				
RL	OncoGene 17:2485-2493(1998).				
CC	- FUNCTION: NOT KNOWN				
CC	- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).				
CC	- TISSUE SPECIFICITY: UBIQUITOUS. EXPRESSED AT HIGH LEVEL IN TESTIS				
CC	AND CNS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF072755; AAC26001.1;				
DR	EMBL; AF093756; AAC78843.1;				
DR	EMBL; AF016398; AAC79938.1;				
DR	MGI; MGI:1316736; Menl.				
KW	Nuclear protein.				
FT	CONFLICT 457 457 M -> I (IN REF. 2).				
FT	CONFLICT 466 466 G -> E (IN REF. 2).				
FT	CONFLICT 512 512 L -> S (IN REF. 2).				
FT	SEQUENCE 611 AA; 67472 MW; C3FB3A5695244E9 CRC64;				
QY	Query Match	95.7%;	Score 3049.5;	DB 1;	Length 611;
Db	Best Local Similarity	96.1%;	Pred. No. 8.6e-191;		
QY	Matches 587;	Conservative	6;	Mismatches	17;
Db				Indels	1;
QY	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVEHFHFLAVNRVPTNPVE	60		
Db	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDVLVLSVLGVEHFHFLAVNRVPTNPVE	60		
QY	61	LTFQSPADPPGGLTYFFVADLSIIAALYARETAIRGAVDLSLYPREGVSSRELKVK	120		
Db	61	LTFQSPADPPGGLTYFFVADLSIIAALYARETAIRGAVDLSLYPREGVSSRELKVK	120		
QY	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITTGTLKDSGGVAFVAVGACQALGLRDVHLASE	180		
Db	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITTGTLKDSGGVAFVAVGACQALGLRDVHLASE	180		
QY	181	HAWVFGPGEQTAETVTHWKGNEDEDRGQTVNAGVAERSWLYLKGSMRCRDKMEVAFV	240		
Db	181	HAWVFGPGEQTAETVTHWKGNEDEDRGQTVNAGVAERSWLYLKGSMRCRDKMEVAFV	240		
QY	241	CAINPSIDLHTDLSLELLOQLKLLWLLYDLGLHLERYPMALGNLADLLELEPTGPRDPLT	300		
Db	241	CAINPSIDLHTDLSLELLOQLKLLWLLYDLGLHLERYPMALGNLADLLELEPTGPRDPLT	300		
QY	301	LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCREDEEI	360		
Db	301	LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCREDEEI	360		
QY	361	YKEFFFEVANDVIPNLLKEASLLBAGERPGEQSGQTSQSALQDPCEFAHLRFDGI	420		
Db	361	YKEFFFEVANDVIPNLLKEASLLBAGERPGEQSGQTSQSALQDPCEFAHLRFDGI	420		
QY	421	CKWEEGSPVLHVGMATFLVQSLGRREGQVRQKRVIRSREAAEAEPGEEAEGRR	480		
Db	421	CKWEEGSPVLHVGMATFLVQSLGRREGQVRQKRVIRSREAAEAEPGEEAEGRR	480		
QY	481	RGPRRSKPEPPPPKPKALDKGLTGQGVSGPPPKPPGTVAGTARGPEGSTAQVPA	539		
Db	481	RGPRRSKPEPPPPKPKALDKGLTGQGVSGPPPKPPGTVAGTARGPEGSTAQVPA	540		

GN BRF3-BERF4.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
ON NCBI_TaxID=10377;
RX SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Sequin C.,
RA Tuffnell P.S., Barrell B.G.:
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RT Nature 310:207-211(1984).
RN [2]
RN CHARACTERIZATION.
RX MEDLINE=88155772; PubMed=2831394;
RX Petti L., Sample J., Wang F., Kieff E.:
RA "A fifth Epstein-Barr virus nuclear protein (EBNA3C) is expressed in
RA latently infected growth-transformed lymphocytes.";
RL J. Virol. 62:1330-1338(1988).
RN [3]
RN SUBCELLULAR LOCATION
RX MEDLINE=90266473; PubMed=2161150;
RA Petti L., Sample C., Kieff E.:
RT "Subnuclear localization and phosphorylation of Epstein-Barr virus
RT latent infection nuclear proteins.";
RL Virology 176:563-574(1990).
CC -!- FUNCTION: INVOLVED IN LATENT CYCLE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -!- SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, 5, AND 6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; V01555; CAA24859.1; -
DR PIR; A03766; Q0B25.
KW Nuclear protein; Repeat.
FT DOMAIN 74 80 POLY-ARG.
FT DOMAIN 551 610 10 X 5 AA TANDEM REPEATS.
FT DOMAIN 741 779 3 X 13 AA TANDEM REPEATS.
SQ SEQUENCE 992 AA; 109129 MW; 39BEAB9BC515BD84 CRC64;
Query Match 3.9%; Score 124.5; DB 1; Length 992;
Best Local Similarity 20.1%; Pred. No. 1.2;
Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;
QY 137 RAHQSLFFSIQTKLSSGVAFVAVGACQALGRDVH-----LALSEDH 181
DB 121 QSHLQALSNLISGLDTHILCFVMAARQR--LQDIRGRLVAEGVGWRHLLTSPSQ 178
QY 182 AWVFGPNQEQTAEV-----TWIGKGNEDRRGQTVAAGV----- 215
DB 179 SW---PMGYRTATLTPVNPVRGADSLMTATFGCQAARTLNTFSVTWTPPHAGP 234
QY 216 -----AERSWLYKSGVMCRDKMEVAFVCAINPSIDLHDSLELQLOQKLLWLY 268
DB 235 REQERYAREAEVFLRGKWRRIYDYLIELCG-----SLHHIWQNLQTEENLLDFVR 289
QY 269 DLGHLERYPMALGNLADLELETPGPRDPLTLYHKGIASAKTY-----RDEHIYPMY 323
DB 290 FMGVMS-----SCNNPAVNWFHKTIGNFKPYYPWNPAPNEN----- 326
QY 324 LAGYHCNRNRVREALQAWADTATVIQDNYCYREDEEYKEFFEVANDVIPNLLKEA---- 379
DB 327 --PYHAR-RGIKE-----HVIQNAFRKAIQIG 350
QY 380 ASLLEAGEERPGSQSGTQS-----QGSALQ-----DPECFAHLRLFYDGICKWEESPT 429

DB 351 LSLMTAGGEPGRGDATSETSSDEDTGRQSDVELESDD-----LPYIDP--NMEPVQQR 403
QY 430 PVLHVG-----WATFLVQSLGRGQVQRKVRIVSREAEAAAEAEPEWGEAREG 478
DB 404 PYMEFSVRPAKKRKLPLWPTPKTHPVKRTNVKTSR-----SDKAEAGSTPE----- 450
QY 479 RRRGPRRES---KP-----EPPPPKKP----- 498
DB 451 -RPGSEQSSVTVEPAHTPVEPMVILHQPVPVKKPVPVKKPTPPRRRRGACVYVDDD 509
QY 499 ---ALDKGLGTQCGAVSGP-----PRKPPGTAGTARGPE--GGSTA 535
DB 510 VIEVIDVETTESSSVSQPNKPHRKHQDGFQSRGRKRAAPPTVSPSDTGPVAVGPPAA 569
QY 536 QVPA---PAASPPPEGP 549
DB 570 GPAAAGPPAAGPPAAGP 586
RESULT 7
PRP1_HUMAN STANDARD; PRT; 331 AA.
AC P04280;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CP5)
DE [CONTAINS: BASIC PEPTIDE IB-6; PEPTIDE P-H].
GN PRB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=85289325; PubMed=2993301;
RA Maeda N., Kim H.-S., Azen E.A., Smithies O.:
RT "Differential RNA splicing and post-translational cleavages in the
RT human salivary proline-rich protein gene system.";
RL J. Biol. Chem. 260:11123-11130(1985).
RN [2]
SEQUENCE OF 214-331.
RX MEDLINE=86243355; PubMed=3521730;
RA Kauffman D., Hofmann T., Bennick A., Keller P.:
RT "Basic proline-rich proteins from human parotid saliva: complete
RT covalent structures of proteins IB-1 and IB-6.";
RL Biochemistry 25:2387-2392(1986).
RN [3]
SEQUENCE OF 276-331.
RX MEDLINE=84161824; PubMed=6671974;
RA Saitoh E., Isemura S., Sanada K.:
RT "Further fractionation of basic proline-rich peptides from human
RT parotid saliva and complete amino acid sequence of basic proline-rich
RT peptide P-H.";
RL J. Biochem. 94:1991-1997(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; K03204; AAA60185.1; -
DR EMBL; K03205; AAA60186.1; -
DR EMBL; K03206; AAA60187.1; -
DR PIR; A03291; PIHUB6.
DR PIR; C25372; C25372.
DR MIM; 168730; -
KW Repeat; Parotid gland; Multigene family; Saliva; Signal.
FT SIGNAL 1 16

```
FT CHAIN 214 331 PEPTIDE IB-6.
FT CHAIN 276 331 PEPTIDE P-H.
FT VARIAT 106 238 MISSING (IN CLONE CP-4).
FT VARIAT 106 258 /FTIG=VAR_005561.
FT VARIAT 106 276 /FTIG=VAR_005562.
FT CONFLICT 276 A -> S (IN REF. 2 AND 3).
SQ SEQUENCE 331 AA: 32596 MW: 34818F8EBA39751 CRC64;

Query Match 3.9%; Score 123; DB 1; Length 331;
Best Local Similarity 38.4%; Pred. NO. 0.35;
Matches 33; Conservative 10; Mismatches 31; Indels 12; Gaps 5;

QY 469 EPMGEAREGRGRRPRE--SKPEPPPPKPKALDKGLGTGGQGVAPPPKPPQTVAGTA 526
Db 46 KPGQPPPPKPGQPPGKNGKPGPPPGKP---QG-PPQGDKSRSPSPKPGKPGPP 101
QY 527 RGPEGGSTAQVPAP-----AASPPPEG 548
Db 102 --PGGNQPGQPPPPKPGKPGPPPG 125

RESULT 8
SEPA_EMENI STANDARD; PRT: 1790 AA.
AC P78621;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOKINESIS PROTEIN SEPA (FHL/2 PROTEIN).
GN SEPA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP MEDLINE=97361829; PubMed=9218790;
RA Harris S.D., Hamer L., Sharpless K.E., Hamer J.E.;
RT "The Aspergillus nidulans sepa gene encodes an Fhl/2 protein involved
in cytokinesis and the maintenance of cellular polarity.";
RL EMBO J. 16:3474-3483(1997).
CC -1- FUNCTION: INVOLVED IN CYTOKINESIS.
CC -1- DOMAIN: DFRS ARE REGULATED BY INTRAMOLECULAR GBD-DAD BINDING WHERE
RHO-GTP ACTIVATES THE DFRS BY DISRUPTING THE GBD-DAD INTERACTION
(BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 GTPASE-BINDING DOMAIN (GBD).
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 1 (FHL) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 2 (FH2) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOLOG 3 (FH3) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 DRF AUTOREGULATORY DOMAIN (DAD).
CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOLOG FAMILY. BN11
SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
FT DOMAIN 1608 1611 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 1790 AA: 198628 MW: 0AB355EC1884D05F CRC64;

Query Match 3.8%; Score 122.5; DB 1; Length 1790;
Best Local Similarity 21.4%; Pred. NO. 3.3;
Matches 110; Conservative 47; Mismatches 187; Indels 171; Gaps 24;

QY 74 GLTYFPVADLSIIAALYAREFTAQIRGAVDLSLYPREGVSSRELVLKKVSDV IWNLSKSY 133
Db 680 GLRMYQLVDAML-----SYVAMDRRLPDLDL--RQG-----LTFVQSLL----- 717
QY 134 FKDRHQISLFSFITGKLDSSGVAFVAGCAQALGRDVLHALSEDHAWVVFPGNG--- 190
Db 718 --DRLHTDAEARRAYDESLEARGIAEA-----ALAEKD-----EMKAQVELCADGLVR 763
QY 191 -----EQTAETVTHCKGNEDRRGTVNAGVAERSWLYLKGSVMCRDKMEVAFWCAI 243
Db 764 KLOQIEEQTGIIELQSRQNE-----MLKAEIADYOR----- 795
QY 244 NPSIDLHTDSLELLOOLKLLWLLY-----DLGHLERYPMALGNLADLEEL 289
Db 796 -----LRAQELQRNELETRELYLMRDAQDIAASNAKSNWGEAETDPAHMRGILDRKEL 850
QY 290 EPTPGRPDLTLYHKGIASAKTYRDEHIYPMYLAGYH---CRNRNVREALQAWADTAT 346
Db 851 -----LTRLKQLERTKQFKLEG-----KVMGQHPDSRLRELREQMDGDAGP-- 894
QY 347 VIQDYNCREDEEYKEFFEVDVNPNL-LKEAASLLEAGEERPGEOSQGTQSQGALQ 405
Db 895 -----REAFE--EQARLNLSLNPVGSYVRKTYIQGMEDTATEELGQT-- 935
QY 406 DPCEFAHLRFYDGICKWEBSGTPVLHVGMATFLVQSLGRFEGOVROKRVISR----- 460
Db 936 DDEVYAKARLVD-----LHRPRMDPEQATGLL-----GEIAKVPKIDADDAKD 980
QY 461 EAEAAAEAEPEWGEAREGRRRG-----PRRESKPEPPPPKPKALDKGLGTGGQAVSG 513
Db 981 EGRPTSEQPAEGAATKGDQGVDDTVAVDKATAAP--PPPPPPPAIPIGI---SGAAPP 1035
QY 514 PPRKPPCTVAGTARGPEGSGTAQVPAPAAASPPPEG 548
Db 1036 PPPPPPPP-----PPGAGAAPPPIPPPPPPPPPG 1063

RESULT 9
PRP2_RAT STANDARD; PRT: 172 AA.
AC P10164;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACIDIC PROLINE-RICH PROTEIN PRP25 PRECURSOR (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RX MEDLINE=86033799; PubMed=3840480;
RA Clements S., Mehansho H., Carlson D.M.;
RT "Novel multigene families encoding highly repetitive peptide
sequences. Sequence analyses of rat and mouse proline-rich protein
cDNAs.";
RL J. Biol. Chem. 260:13471-13477(1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

or send an email to license@isb-sib.ch).

EMBL: M58378; AAC41930.1; JOINED.
EMBL: M58321; AAC41930.1; JOINED.
EMBL: M58341; AAC41930.1; JOINED.
EMBL: M58351; AAC41930.1; JOINED.
EMBL: M58353; AAC41930.1; JOINED.
EMBL: M58359; AAC41930.1; JOINED.
EMBL: M58371; AAC41930.1; JOINED.

5

[illegible]

EMBL; M58373; AAC41931.1; JOINED.
EMBL; M58374; AAC41931.1; JOINED.
EMBL; M58375; AAC41931.1; JOINED.
EMBL; M58376; AAC41931.1; JOINED.
EMBL; M58377; AAC41931.1; JOINED.
EMBL; M55301; AAA60608.1; "-.
EMBL; AL009172; CAA15657.1; "-.
PIR; A35363; A35363.
PIR; B35363; B35363.

MIM: 313440; -.
InterPro: IPR001359; -.
Pfam: PF02078; Synapsin: 1.

PROSITE; PS00412; SNAFESINL1; 1.	A.	28
PROSITE; PS00416; SNAFESINL2; 1.	B (LINKER).	29
Synapse; Phosphorylation; Neuron; Repeat; Actin-binding.	C (ACTIN-BINDING AND SYNAPTIC BINDING).	113
Synapse; Phosphorylation; Neuron; Repeat; Actin-binding.		420
Alternative splicing.		
DOMAIN		1
DOMAIN		29
DOMAIN		113
DOMAIN		420

[illegible]

CONFLICT 138 G \rightarrow E {IN REF. 3).
SEQUENCE 705 AA: 73954 MW: 487831123FF6882F CRC64:

Very Match 3.7%; Score 118; DB 1; Length 705;
1st Local Similarity 20.8%; Pred. No. 2;
Matches 132; Conservative 67; Mismatches 213; Indels 222;
Gaps 34;

lines 132, conservative 67, M

66 SPAPDPGGGLTYFPVADLSI---IAALYARFTAQIRGAVDLSLYPREGGVSSRELV----118

65 APSPGSSGGGGFESSLSNAVKOTTAAAAATFSEOVGGG--SGGAGRGGAASRVLLVIDE 121

119 -----KKVSDVIWNSI--SRSYFKDR---AHIQSI--FSEITGTKIDSSGVAFAVUGA 164

[illegible]

122 PHTDNAKYFGKK IHGGIDIKVEQAEFSDLNLVAHANGCES--VDMEVLRNGV--KVVRs 177
165 COALGLRDVHLIALSDHAWVFGPNEGQTAEVTHHGKGNEDRRGQTVN---AGVAERSWL 221
178 LKP-----DFVLIHQAFSM-----ARNGDYRSLVTGIIQYAGIPSVNSL 216
222 YLKGSYMRCDRKMEVAFVWCVAINPSIDLH-----TDSLELL-----OLOOKLILWL 257

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT THE KIDNEYS,
 CC BUT LEVELS ARE HIGHEST IN THE HEART.
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-16 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF055994; AAC39854.1; -;
 DR EMBL; Y13467; CAA73867.1; -;
 DR EMBL; L40366; AAC41736.1; -;
 DR MIM; 604311; -;
 KW DNA-binding; Transcription regulation.
 FT DOMAIN 604 649 2 X REPEATS OF L-X-X-L-L.
 FT REPEAT 604 608 1.
 FT REPEAT 645 649 2.
 FT DOMAIN 1099 1111 POLY-SER.
 FT DOMAIN 1269 1279 POLY-SER.
 FT CONFLICT 86 86 R -> G (IN REF. 2).
 FT CONFLICT 147 147 F -> S (IN REF. 2).
 FT CONFLICT 471 472 DS -> GL (IN REF. 2).
 FT CONFLICT 543 545 LAG -> ASS (IN REF. 2).
 FT CONFLICT 563 563 P -> S (IN REF. 2).
 FT CONFLICT 573 573 T -> A (IN REF. 2).
 FT CONFLICT 573 573 S -> F (IN REF. 3).
 FT CONFLICT 702 708 MISSING (IN REF. 3).
 FT CONFLICT 721 721 K -> N (IN REF. 2).
 FT CONFLICT 1388 1388 S -> G (IN REF. 2).
 FT SEQUENCE 1581 AA; 168517 MW; DD726AF6885AFF57 CRC64;

Query Match 3.7%; Score 117.5; DB 1; Length 1581;
 Best Local Similarity 22.4%; Pred. No. 5.9;
 Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDYNVCRDEIYKEFEVANDVLPNLLKEAASLLEAGEERPEQSQ-----395
 DB 835 YTDPADLIADAAGSPSSDPTNHFHGDGVDFNPDLN-NSQSGGFEEDFSSQSGDND 893
 QY 396 ---GTQSQ-----GSAQDQPECFALLRF 416
 DB 894 DFKGFASQALNTLGVPLMGDNGETKFKGNQADTVDFSIISVACKALAPADLMEH----949
 QY 417 YDGICKWEGSPTPVLHVGNATFLVOSLGRFEGQVQKRVIVSRREAAEAEPWGEAR 476
 DB 950 -----HSGSQQLITTC-----DLGKEKTQKRVK---EGNGTSTSLSGP-GLDSK 991
 QY 477 EGRG-RGRRESKPEPPPKPALDKGLGTGQAVSGPPRPPTVAGTARCEGGSTA 535
 DB 992 PKRSTRTPNSDKSKDKPKPKKADTEGKSPSHSS--SNRPFTPP--TSTGKSPSGSAGRS 1049
 QY 536 QVPAPAAAPP-----PEGVPLTFQSEKMGKMKELLVATKINSIAIK-----LQL 579
 DB 1050 QTPPGVATPPPKITIQIPKGVWV---CKPSSHQSYTSGSVSSGSKSHSHSSSSS 1106
 QY 580 TACSQVMKKQKVPSTPDTLS 601
 DB 1107 SASTSGKMKSSKSGSSSKLS 1128

RESULT 13
 MACS_MOUSE
 ID MACS_MOUSE STANDARD; PRT; 308 AA.
 AC P26645;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).
 GN MACS.

OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=Macrophage;
 RX MEDLINE=91172836; PubMed=2006186;
 RA Seykora J.T., Ravetch J.V., Aderem A.;
 RT "Cloning and molecular characterization of the murine macrophage '68-
 RT kda' protein kinase C substrate and its regulation by bacterial
 RT lipopolysaccharide.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2505-2509(1991).
 RN [2]
 RN RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=91330872; PubMed=1868832;
 RA Brooks S.F., Herget T., Erusalimsky J.D., Rozengurt E.;
 RT "Protein kinase C activation potentially down-regulates the expression
 RT of its major substrate, 80K, in Swiss 3T3 cells.";
 RL EMBO J. 10:2497-2505(1991).
 RN [3]
 RN RP PARTIAL SEQUENCE.
 RC STRAIN=SWISS; TISSUE=Fibroblast;
 RX MEDLINE=90346162; PubMed=2384168;
 RA Brooks S.F., Erusalimsky J.D., Totty N.F., Rozengurt E.;
 RT "Purification and internal amino acid sequence of the 80 kDa protein
 RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with
 RT substrates from brain.";
 RL FEBS Lett. 268:291-295(1990).
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
 CC -!- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND
 CC VERY LOW LEVELS IN LIVER.
 CC -!- INDUCTION: BY LIPOPOLYSACCHARIDE.
 CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO THE MARCKS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M60474; AAA39491.1; -;
 DR PIR; A39169; A39169.
 DR PIR; S16519; S16519.
 DR MGD; MGI:96907; Macs.
 DR InterPro; IPR002101; -;
 DR Pfam; PF02063; MARCKS; 1.
 DR PRINTS; PR00963; MARCKS.
 DR PROSITE; PS00826; MARCKS_1; 1.
 DR PROSITE; PS00827; MARCKS_2; 1.
 DR PROSITE; PS00827; MARCKS_2; 1.
 KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
 KW Membrane.
 FT INIT MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT DOMAIN 144 168 CALMODULIN-BINDING (PSD).
 FT MOD_RES 151 151 PHOSPHORYLATION (BY PKC). (BY SIMILARITY).
 FT MOD_RES 155 155 PHOSPHORYLATION (BY PKC). (BY SIMILARITY).
 FT MOD_RES 159 159 PHOSPHORYLATION (BY PKC). (BY SIMILARITY).
 FT MOD_RES 162 162 PHOSPHORYLATION (BY PKC). (BY SIMILARITY).
 FT CONFLICT 95 97 AGA -> TGT (IN REF. 2).
 FT SEQUENCE 308 AA; 29530 MW; FB5313B913701C5C CRC64;

Query Match 3.7%; Score 117; DB 1; Length 308;
 Best Local Similarity 26.3%; Pred. No. 0.79;
 Matches 54; Conservative 21; Mismatches 86; Indels 44; Gaps 8;

592 PPGPAGPIRQASQAG-----PGPRTGPTTQQPRPSCP 524

RESULT 15

ADDB_HUMAN STANDARD; PRT; 726 AA.

AC P35612; Q13482; Q16412;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE BETA ADDUCIN (ERYTHROCYTE ADDUCIN BETA SUBUNIT).

GN AD2 OR ADDB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

11

SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.

RP TISSUE-Reticulocytes;

RX MEDLINE=92011907; PubMed=1840603;

RA Joshi R.L., Gilligan D.M., Otto E., McLaughlin T., Bennett V.D.;

RT "Primary structure and domain organization of human alpha and beta adducin";

RT J. Cell Biol. 115:665-675(1991).

12

SEQUENCE OF 186-726 FROM N.A. (ISOFORM 3).

RA Sun H., Abbott A.;

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

13

SEQUENCE OF 332-726 FROM N.A. (ISOFORM 3).

RP TISSUE=Bone marrow;

RA Sinard J.H., Stewart G.W., Argent A.C., Gilligan D.M., Morrow J.S.;

RT "A novel isoform of beta adducin utilizes an alternatively spliced exon near the C-terminus";

RL Mol. Biol. Cell 6:269-269(1995).

14

ALTERNATIVE SPLICING.

RP MEDLINE=96144294; PubMed=8566798;

RA Tsimenetsky S., Devescovi G., Tripodi G., Muro A., Bianchi G., Colombi M., Moro L., Barlati S., Tuteja R., Baralle F.E.;

RT "Genomic organisation and chromosomal localisation of the gene encoding human beta adducin";

RL Gene 167:313-316(1995).

15

PHOSPHORYLATION SITES, AND PARTIAL SEQUENCE.

RP MEDLINE=96411719; PubMed=8810272;

RA Matsuoka Y., Hughes C.A., Bennett V.;

RT "Adducin phosphorylation. Definition of the calmodulin-binding domain and sites of phosphorylation by protein kinases A and C.";

RL J. Biol. Chem. 271:25157-25166(1996).

16

FUNCTION: MEMBRANE-CYTOSKELETON-ASSOCIATED PROTEIN THAT PROMOTES THE ASSEMBLY OF THE SPECTRIN-ACTIN NETWORK. BINDS TO CALMODULIN. CALMODULIN BINDS PREFERENTIALLY TO THE BETA SUBUNIT.

17

SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.

18

ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/ADDUCIN 63 AND 3/BETA-4; ARE PRODUCED BY ALTERNATIVE SPLICING.

19

TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN, SPLEEN AND K562 CELLS.

20

DOMAIN: EACH SUBUNIT IS COMPRISED OF THREE REGIONS: A NH2-TERMINAL PROTEASE-RESISTANT GLOBULAR HEAD REGION, A SHORT CONNECTING SUBDOMAIN, AND A PROTEASE-SENSITIVE TAIL REGION.

21

PTM: THE N-TERMINUS IS BLOCKED.

22

SIMILARITY: BELONGS TO THE ADDUCIN FAMILY.

23

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X58199; CAA41176.1; -

```

DR EMBL; AC005234; AAd12716.1; -.
DR EMBL; U43959; AAA86421.1; -.
DR EMBL; S81079; AAd14349.1; -.
DR EMBL; S81077; AAd14349.1; JOINED.
DR PIR; S18208; S18208.
DR MIM; 102681;
DR InterPro; IPR001303; -.
DR Pfam; PF00536; Aldolase_II; 1.
KW Cytoskeleton; Membrane; Calmodulin-binding; Phosphorylation;
KW Alternative splicing.
FT DOMAIN 425 444
FT MOD_RES 704 721
FT MOD_RES 55 55
FT MOD_RES 703 703
FT MOD_RES 713 713
FT VARSPPLIC 532 560
FT FT
FT VARSPPLIC 561 726
FT VARSPPLIC 581 726
FT FT
FT LEETGKTETSKAATTEPTQPEGVVVNGREEEQTAEEIL
FT SKGLSQMTTSADTDVDTSKDTESVTSGPMSPGSPKSPS
FT KKKKTKTPSFLAKKKKKEVES -> ETGOEREPPGSGPAV
FT CEPESVALHWSNILEKKLPQKSLAHLQSLHILLQCRAR
FT RRRQRRL (IN ISOFORM 3).
FT FT
SQ SEQUENCE 726 AA; 80854 MW; B07F730D929DBA4 CRC64;

Query Match 3.6%; Score 116; DB 1; Length 726;
Best Local Similarity 20.3%; Pred. No. 2.8;
Matches 8; Conservative 48; Mismatches 134; Indels 140; Gaps 17;

Qy 245 PSIDLHTDSLELLOOLKMLLLYDLGLHRYPMALGNLADLEELETPGPRDPLTLYHK 304
Db 417 PALRQHAQK----QOKEKTRLNTPTNYLR-----VNVAD--EVQRSMGSPRKTMMK 464
Qy 305 GIASAKTYRDEHIYPWYLAGYHCRNRNVRALQANADTATVTDNYCREDEIYKEF 364
Db 465 ADEVEKS-----SSGMPRIRENQFVPLTYDPQEVLEMKNRIKE-----504
Qy 365 FEVANDVIPNLLKEAASLLEAGEEPGEGSGTGSQ-----GSALODPECFAHLRFRYD 418
Db 505 -----QNRQDVKSAGPOSQALLASVIAKSRSPSTESQMSKGD 542
Qy 419 GICK--WEEGSPTVLHVGNATFLVQSLGRPEGQVRQKRVIRSBREAFAAEPPWGEAR 476
Db 543 EDTKDDSEETVPNP-----PSQLTDQEELEYKKEVERKKLELDGEKETA-----586
Qy 477 BGRRRGPRRESKPEPPPPPKPALDKGCTQGAGVGGPPRPPGCTVAGTARCP-----529
Db 587 -----PEEPGSPAKSA-----PASPVGQSPAKEET-KSPLVSPSKS 621
Qy 530 --EGGSAQVAPAPASPP-----PEGPVLTFQSEKM-----KGMKELLVA--TKINSSA 574
Db 622 LEETGKTETSKAATTEPTQPEGVVVNGREEEQTAEEILSKLSQMTTSADTDVDTSK 681
Qy 575 IKQLQTAQSQV-----QMKKQKVSTPSDYTLNFLKQRK 608
Db 682 DKTESVTSGPMSPGSPKSPSKKKKFRTP-----SFLKSKK 720

Search completed: October 12, 2001, 18:22:11
Job time: 199 sec

```

Search completed: October 12, 2001, 18:22:11
Job time: 199 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:14:32 ; Search time 19.39 Seconds
(without alignments)
2396.414 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLPLRSIDVV.....KVSTPSDYTLFLKQRKGL 610
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_58:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126.5	4.0	813	2 T02672	hypothetical prote
2	125.5	3.9	1025	2 S34839	collagen alpha 1(V
3	124.5	3.9	872	2 S33015	hypothetical prote
4	124.5	3.9	992	2 A31666	hypothetical prote
5	124	3.9	392	1 PIHUB6	salivary proline-r
6	123.5	3.9	1560	2 T02885	peroxisome prolif
7	123.5	3.9	3436	2 S55659	tegumentary protei
8	122.5	3.8	980	2 S54986	regulatory protein
9	119	3.7	172	2 B25149	proline-rich prote
10	119	3.7	705	2 A33363	synapsin I splice
11	118.5	3.7	117	2 D40750	proline-rich prote
12	118	3.7	7962	2 I38346	elastic titin - hu
13	117.5	3.7	251	1 PIHUPF	salivary proline-r
14	117	3.7	309	2 A39169	myristylated alani
15	117	3.7	552	2 T08148	proline-rich myros
16	116.5	3.7	223	2 A42817	proline-rich prote
17	116.5	3.7	2109	2 I38414	transcription fact
18	116	3.6	704	2 A30411	synapsin Ia - rat
19	116	3.6	726	2 S18208	rabphilin-3A-inter
20	116	3.6	2774	2 A43359	microtubule-associ
21	115	3.6	128	2 D38355	basic proline-rich
22	115	3.6	1237	2 T45070	protein kinase hom
23	114.5	3.6	479	2 S10127	muscarinic acetyl
24	114.5	3.6	637	2 T04552	hypothetical prote
25	114	3.6	492	2 A41907	methyl-Cpg-binding
26	113.5	3.6	1838	1 CGHUIV	collagen alpha 1(V
27	113	3.5	577	2 T00369	hypothetical prote
28	113	3.5	706	2 E30411	synapsin Ia - bovi
29	113	3.5	1486	1 B40333	collagen alpha 1(I

30	112.5	3.5	310	1 PIHUSD	salivary proline-r
31	112	3.5	188	2 JH0481	basic proline-rich
32	111.5	3.5	2453	2 S60254	nuclear receptor c
33	111	3.5	686	2 A38235	microtubule-associ
34	111	3.5	695	2 T36007	probable transketo
35	111	3.5	744	2 T35192	probable ABC trans
36	111	3.5	1747	1 A45974	collagen alpha 1(X
37	111	3.5	1857	2 S31212	collagen alpha 1(X
38	111	3.5	1888	2 S78476	collagen alpha 1(X
39	110.5	3.5	441	2 T42694	hypothetical prote
40	110	3.5	479	2 S33776	muscarinic acetyl
41	110	3.5	754	2 A56619	female sterile hom
42	110	3.5	858	2 S15762	neurofilament trip
43	110	3.5	1419	2 A41182	collagen alpha 1(I
44	110	3.5	1494	2 T14355	protein-tyrosine-p
45	110	3.5	1844	2 T51890	related to Nup98-N

ALIGNMENTS

RESULT 1
T02672
hypothetical protein R31449_3 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 31-Dec-2000
C:Accession: T02672
R:Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.
J.; Danganan, L.; Poundstone, P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.;
P.; Quan, G.; Kronmiller, B.; Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.
submitted to the EMBL Data Library, June 1998
A:Authors: Trong, S.; Kobayashi, A.; Olsen, A.S.; Carrano, A.V.
A:Description: Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a ser
A:Reference number: Z14696
A:Accession: T02672
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-813 <L>AM>
A:Cross-references: EMBL:AC005175; NID:g3253116; PID:g3253120
C:Genetics:
A:Map position: 19
A:Introns: 93/2; 125/2; 174/1; 228/2; 266/3; 331/1; 372/1; 474/3; 637/1; 678/3
A:Note: R31449_3

Query Match	4.0%;	Score 126.5;	DB 2;	Length 813;
Best Local Similarity	20.6%;	Pred. No. 0.81;		
Matches	94;	Conservative	74;	Mismatches 172; Indels 117; Gaps 21;
QY	204	EDRRGOTVNAVGAERSWLYL-KGSYMRCDKMEVAFMVCAINPSIDLHTDSLELLQLOQK 262		
Db	103	EEKELAKERNKRIQEDNRLQKQVQLRLERREKA-----MREQELEMLOREKE 151		
QY	263	L-----LW-LLYDLGLHLYPYMALGNLADLELEPTGRRDPDLTYLHKGIASAKTYRDEH 317		
Db	152	AHFKEWEEQEDNFHLQQAALR-----SKIRIDGRAPKPIDLLAKYISAEDDLAVEM 204		
QY	318	IYPYMYLAGYHCNRNVREALQAWADATATVIQDYNCRDE-----EIVKEFEVANDVI 372		
Db	205	HEPYTFNLGL-----TVADMEDLLDIQVYMELEQGNKADFWRDMTTTIDEI 252		
QY	373	PNLLKEAASLLEAGEERPGEGSQGSQGS-----ALQDPECFAHLRLFYDGI-CKWEES 427		
Db	253	SKLRK-----LEASGKCPGERREGVNASVSSDVQSVFKGKTYNQLOQVIFOGIEGKIRAGG 307		
QY	428	PTPVLHVGNATFLVQSL-----GRFEGQVQRKVRIVSREAEAAAE-----E 469		
Db	308	PN--LDMGYWESLLQQLRAHMAARLRERHQDVLRLKLYKLKQE-QGVSEPLFPIKQE 364		
QY	470	PGEEAREGRRRGPRRESPEEPKPKALDKGLGTGGQGVSG-PPRKPPGTAVGTARG 528		
Db	365	P-----QSPSRSLPEDAAPTTPPGPSSEG-GPAEAVDGTPTGDDGDEGEG 413		

Qy 529 PEGSTAQVPAPAAAPPEGVLFTQSEKMKMKE-----LLVATKINSATKIQ 578
Db 414 -----EGEAVLMEEDLIQSLDDYDAGRYSPRLTAHELPLDAHLE 455
Qy 579 LTAQSQ-VQMKKQKVPSTPDTLS-----FLKQRKGL 610
Db 456 PDELQRLQLSROOLQVTGDAESAEDIFFRRAKEGM 492

RESULT 2
S34839
collagen alpha 1(VI) chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C:Accession: S34839; S31405; S32603; S31403; S32003
R:Bonaldi, P.; Piccolo, S.; Marvulli, D.; Volpin, D.; Marigo, V.; Bressan, G.M.
Matrix 13, 223-233, 1993
A:Title: Murine alpha-1(VI) collagen chain. Complete amino acid sequence and identification submitted to the EMBL Data Library, May 1992
A:Reference number: S34839; MUID:93316904
A:Accession: S34839
A:Molecule type: mRNA
A:Residues: 1-1025 <BOW>
A:Cross-references: EMBL:X56405; NID:g50478; PIDN:CAA47032.1; PID:g50479
R:Bonaldi, P.; Piccolo, S.; Marvulli, D.; Volpin, D.; Bressan, G.M.
submitted to the EMBL Data Library, May 1992
A:Reference number: S31403
A:Accession: S31405
A:Molecule type: DNA
A:Residues: 1-31 <BOW>
A:Cross-references: EMBL:X56406; NID:g51055; PIDN:CAA47033.1; PID:g51056
R:Zhang, R.Z.; Pan, T.C.; Timpi, R.; Chu, M.L.
Biochem. J. 291, 787-792, 1993
A:Title: Cloning and sequence analysis of cDNAs encoding the alpha-1, alpha-2 and alpha-3 chains of human collagen alpha 1(VI) chain. Complete amino acid sequence and identification submitted to the EMBL Data Library, March 1988
A:Reference number: S32603; MUID:93256888
A:Accession: S32603
A:Molecule type: mRNA
A:Residues: 442-673, 'TL', 676-708, 'A', 710-942, 944-959, 'R', 961-1025 <ZHA>
A:Cross-references: EMBL:218271; NID:g57955; PIDN:CAA79152.1; PID:g57956
C:Genetics:
A:Gene: COL6A1
C:Superfamily: collagen alpha 1(VI) chain; von Willebrand factor type A repeat homology
C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; heterotrimer; F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1025/Product: collagen alpha 1(VI) chain #status predicted <MAT>
F;34-212/Domain: von Willebrand factor type A repeat homology <WVAL>

Query Match 3.9%; Score 125.5; DB 2; Length 1025;
Best Local Similarity 21.4%; Pred. No. 1.3;
Matches 94; Conservative 36; Mismatches 135; Indels 175; Gaps 23;

Qy 223 LKGYMRCRDKM-----EV-----AFMWCA 242
Db 72 LRDRYRCDRNLVWNGALHYSDEVEIRGLTRMPSGRDELKASVDKVGKTYTDCA 131
Qy 243 INPSIDLHSDLELL-----QLQKLLMLLDLGH-LERYPMALGNLAD-LEELEPTGPR 296
Db 132 IKKGL-----ELLIGSHLKENKYLIVTDGHPLEGYKEPCGLEDVANEAK----- 179
Qy 297 DPLTYHKGASATYRDEHIYPM-YLAGYHCRNRNVREALQAWDATATVIQDYNCR 355
Db 180 -----HLGIKVFSAVTPDHLPRLSLIATDHYRRNFTAA-----DWGHSR 221
Qy 356 EDEIYKEFFEVANDVIPNLLKEASLEA-----GEE-- 388
Db 222 DAEEVISQITDITVDMIKNNVQCCSFECQAAAGPPGPRCDPCYEGERGKPLGEGKE 281
Qy 389 -----RCEQS-----QGTQSQ-GS-----ALQDPCF--AHLIRFYDGI----- 420
Db 282 AGDPGRPGDLGPVGYQGMKGKSGRKGSGRKGKRGIDGDMKGTGYPGL 341
Qy 421 --CKWE-----EGSPTPVLHVGNATLVQSLGRFEGQVRQKRVISREAEAEPEWG 472

Db 342 PGCKSGPGFDGIQPGPKGDAG-----AFCKMKGKGGAGAD-G 379
Qy 473 EEARERRRRGRPRRESKPEEPPEPP--KKPALDKGLGTGGAVSGPPRPKPGTVACTA-RGP 529
Db 380 EAGRPGNSGPDGEGDGEFGPGGKEGAGDEGNAGPDGA-----PGERGGPGERGP 431
Qy 530 EGGSTAQVPAPAAAPPPPEGP 549
Db 432 RG--TPGVRGPRGDPGEAGP 449

RESULT 3
S33015
hypothetical protein - human herpesvirus 4
C:Species: human herpesvirus 4, Epstein-Barr virus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Aug-1997
C:Accession: S33015
R:Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A:Reference number: S32973
A:Accession: S33015
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-872 <FAR>
A:Cross-references: EMBL:V01555
C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3C

Query Match 3.9%; Score 124.5; DB 2; Length 872;
Best Local Similarity 20.1%; Pred. No. 1.2;
Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;

Qy 137 RAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGLRDVH-----LALSDEH 181
Db 1 QSHLQALSNLILDSGLDTQHILCFVMAARQR--LQDIRGGLVAEGGVGRHMLLTSPS 58
Qy 182 AWVFGNGEGQTAEV-----TWHGKGNEDRRGQTVNAGV----- 215
Db 59 SW-----PMGYRTATRLTLPVNRVGADSIMLTATFCQNAARTLNTFSATVTPPIHAGP 114
Qy 216 -----AERSWLYLKGSYRCDRKMVEAFVCAINSIDLHTDSLELLOLQKLLMLLY 268
Db 115 REQRYAREAEVRLGRKWRRYRRIYDLIELCG-----SLHHWQNLLOTEENLLDFVR 169
Qy 269 DLGHLERYPMALGNLADLEELETPGPRDPLTLTHKGIASAKTYV---PDEHIYPYMY 323
Db 170 FMGYS-----SCNNPANYWPHKTIQNFKPYYPNAPNEN----- 206
Qy 324 LAGYHCRNRNVREALQAWDATATVIQDYNYCREDEEYKEFFEVDNDVIPNLLKEA---- 379
Db 207 --PYHAR-RGIKE-----HVIONAFRKAQIQG 230
Qy 380 ASLLEAGEERPGESQGTOS-----QGSALO-----DPECFAHLIRFYDGIKKEGSPT 429
Db 231 LSMIATGEPRGDQATSETSSDEDTGROGSDVELESSODE-----LPYIDP--NNPEVQOR 283
Qy 430 PVLHVG-----WATFLVQSLGRFEGQVRQKRVISREAEAEALPWCGEAREG 478
Db 284 PVMFVSRVPAKKPRKLPWPTKPTHVKRTNVKTSR-----SDKAEAOSTPE----- 330
Qy 479 RRGPRRES-----KP-----EPPPPPKP----- 498
Db 331 -RPGSQSSVTVPEFAHPTVPMPVILHQPVPVKKPVKPTPPPSRRRRKRGACVVYDDD 389
Qy 499 ---ALDKGLGTGGCAVSGP-----PRKPPGTVAGTARGPE--GGSTA 535
Db 390 VIEVIDVETTEDSSVSQPNKPHRKHQDQFQSGRRGRKRAAPTVPSDITGPPAVGPAA 449
Qy 536 QVPA---PAASPPPEGP 549
Db 450 GPAAAGPPAAGPPAAGP 466

RESULT 4

A31666
 hypothetical protein - human herpesvirus 4
 C:Species: human herpesvirus 4, Epstein-Barr virus
 C>Date: 20-Jul-1989 #sequence_revision 20-Jul-1989 #text_change 21-Jul-2000
 C:Accession: A31666; A49034; B49034; C49034; D49034; E49034; F49034
 R:Sawada, K.; Yamamoto, M.; Tabata, T.; Smith, M.; Tanaka, A.; Nonoyama, M.
 Virology 168, 22-30, 1989
 A:Title: Expression of EBNA-3 family in fresh B lymphocytes infected with Epstein-Barr virus
 A:Reference number: A31666; MUID:89085606
 A:Accession: A31666
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-992 <SAW>
 R:Apolloni, A.; Moss, D.; Stumm, R.; Burrows, S.; Suhrbier, A.; Misko, I.; Schmidt, C.;
 Eur. J. Immunol. 22, 183-189, 1992
 A:Title: Sequence variation of cytotoxic T cell epitopes in different isolates of Epstein-Barr virus
 A:Reference number: A49034; MUID:92111623
 A:Contents: nuclear antigen EBNA-6
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 281-300 <APO>
 A:Cross-references: GB:S79197; NID:9242871; PIDN:AAB20975.1; PID:9242872
 A:Experimental source: type A strain B95-8; type A strain IARC/BL74; type A strain IARC/BL74; type A strain IARC/BL74; type A strain IARC/BL74
 A:Note: sequence extracted from NCBI backbone (NCBIN:79200, NCBI:79212)
 C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3C

Query Match 3.9%; Score 124.5; DB 2; Length 992;
 Best Local Similarity 20.1%; Pred. No. 1.4;
 Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;

QY 137 RAHQSLFETITKLDSSGVAFVYVACQALGLRDVH-----LALSEDH 181
 Db 121 QSHLQALNLLDGLDTHQILCFYMAQR--LQDIRGLPLVAEGVGRHWLLTSPSQ 178
 QY 182 AWVYFGNGEQTAEV-----TWHGKGNDRRGQTNAVY----- 215
 Db 179 SW----PMGYRTATRLTPVPRVGVADSIMLTATGCGNAARTLNTSATVTPPHAGP 234
 QY 216 -----AERSWLYKSGWRCRDKMEVAFVCAINPSIDLTFDLSLELQLOKLLWLLY 268
 Db 235 REQERYAREAEVFLRGKQWRYYRITLIELCG-----SLHHWQNLQTEENLLDFVR 289
 QY 269 DLGHLERYPMALGNLADLELETPGRRPDLTLYHKGIASAKTY-----RDGHIYPYMY 323
 Db 290 FGVMS-----SCNNPAVNWFHKTIGNFRPYPNWAPPNEN----- 326
 QY 324 LAGYHCNRRNRREALQAWDATVIQDYNVCREDEEYKEFFEVANDVIPNLKEA---- 379
 Db 327 --PYHAR-RGIKE-----HVIQNAFRKAQIQG 350
 QY 380 ASLLENGERPGSQSGTQS-----QGSALQ-----DPECFAHLRFYDGIKWEGSGPT 429
 Db 351 LSLMATGGEPRGDATSETSSDEDTGQGSVLESSDDE-----LPYIDP--NNEPVPQOR 403
 QY 430 PVLHVG-----WATFLVQSLGREGGVQRVIRVSRERAEAAEAEPGGEAREG 478
 Db 404 PMFVSVRPAKPKRPLPWPTTKHPVKRTNVKTSR-----SDKAEAQSTPE----- 450
 QY 479 RRRGPRRES----KP-----BEPPPPKP----- 498
 Db 451 -RGPSEQSSVTVEPAHPTVPMPVILHQPPVPKPTPPRRRRRGACVYDDDD 509
 QY 499 ---ALDKGLTGCGAVSGP-----PRKPPGTAGTARGPE--GGSTA 535
 Db 510 VIEVIDVETTEDSSVSQPNKPHRKHQDGFQSRGRKRAAPPTVSPSDTGPVAVGPPAA 569
 QY 536 QVPA---PAASPPPEGP 549
 Db 570 GPPAAGPAAGPPAAGP 586

RESULT 5
 PIHUB6
 salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - huma
 N:Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H
 C:Species: Homo sapiens (man)
 C>Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 08-Dec-2000
 C:Accession: B40750; C40750; A0750; C25372; S02127; A03293; A90502; A91974;
 R:Azen, E.A.; Latreille, P.; Niece, R.L.
 Am. J. Hum. Genet. 53, 264-278, 1993
 A:Title: PRB1 gene variants coding for length and null polymorphisms among human sali
 A:Reference number: A40750; MUID:93304421
 A:Accession: B40750
 A:Molecule type: DNA
 A:Residues: 35-392 <AZE>
 A:Cross-references: GB:S62941
 A:Experimental source: subject C.J. (large allele)
 A:Accession: C40750
 A:Molecule type: DNA
 A:Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-39
 A:Cross-references: GB:S62929
 A:Experimental source: subject M.V.O. (large allele)
 A:Accession: A40750
 A:Molecule type: DNA
 A:Residues: 1-183,245-270,'Q',272-392 <AZ3>
 A:Cross-references: GB:S62928
 A:Experimental source: subject C.J. (medium allele)
 A:Note: authors translated the codon CAA for residue 272 as Arg
 R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
 J. Biol. Chem. 260, 11123-11130, 1985
 A:Title: Differential RNA splicing and post-translational cleavages in the human sali
 A:Reference number: A92492; MUID:85289325
 A:Accession: C25372
 A:Molecule type: mRNA
 A:Residues: 1-183,245-392 <MAE>
 A:Cross-references: GB:K03204; NID:gl90485; PIDN:AAA60185.1; PID:gl90486
 A:Note: alternatively splice forms lacking portions of the repeat region were also fo
 R:Lyons, K.M.; Stein, J.H.; Smithies, O.
 Genetics 120, 267-278, 1988
 A:Title: Length polymorphisms in human proline-rich protein genes generated by intrag
 A:Reference number: S02127; MUID:89121440
 A:Accession: S02128
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>
 A:Cross-references: EMBL:X07517
 A:Accession: S02127
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 35-183,245-392 <LY2>
 A:Cross-references: EMBL:X07516
 R:Kaufman, D.; Hofmann, T.; Bennick, A.; Keller, P.
 Biochemistry 25, 2387-2392, 1986
 A:Title: Basic proline-rich proteins from human parotid saliva: complete covalent str
 A:Reference number: A90502; MUID:86243355
 A:Accession: A03293
 A:Molecule type: protein
 A:Residues: 17-38,AP',41-51,92-148,'R',150-152 <KA2>
 A:Note: among nine basic proline-rich peptides isolated from the saliva, this peptide
 A:Accession: A90502
 A:Molecule type: protein
 A:Residues: 275-336,'S',338-392 <KAU>
 R:Saich, E.; Isemura, S.; Sanada, K.
 J. Biochem. 94, 1991-1999, 1983
 A:Title: Further fractionation of basic proline-rich peptides from human parotid sali
 A:Reference number: A91974; MUID:84161824
 A:Contents: P-H
 A:Accession: A91974
 A:Molecule type: protein
 A:Residues: 'S',338-392 <SAI>
 R:Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Van
 Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984
 A:Reference number: A94005; MUID:84298176

Query Match	3.9%	Score 123.5;	DB 2;	Length 1560;
Best Local Similarity	23.2%	Pred. No. 3.1;		
Matches 73;	Conservative	39;	Mismatches 131;	Indels 71;
Matches 13;				Gaps 13;

Query Match	3.9%;	Score 123.5;	DB 2;	Length 3436;
Best Local Similarity	22.3%;	Pred. No. 8.8;		
Matches	79;	Conservative 33;	Mismatches 112;	Indels 131;
Gaps				
Qy	314	RDEHIY---PYMKLACYHC-----	-----RNRN-VREALQAWDATATVIQDYN	352
Db	3052	RDRFVVLDPNNYLLKGSYLGNAHPGDAPPV	ELAIENKNWPAEALPNLQAASPIPKSR	311
Qy	353	YCREDEIYKEFFEVANDVIPNLLKEA----	ASLLEAGEERPGSQSQSQGSAQLDP	407
Db	3112	VAREYKLLTEKYMDLA-----	LKQVFAEFGSVHDDPSDREEEGECEACQAE	316
Qy	408	ECFAHLLRFYDGICKWEE-GSPFPVLHVGMAT	FLVOSLGRFGQVRQKRVIVSRE----	463
Db	3165	E-----EWEREERPPRRLG-----	ERGRREVEVAVAPADPRDYSRPR	320
Qy	464	AARAEPPWG-----	BEAREGRRR-----	GPRR 485
Db	3205	VPPRREPFGSPRSPARDRTLGGGAGSRKEVSE	GRGPRVQLSRSPKPRPAASQVOGPPE	326
Qy	486	E-----SKPEEP-----	PPKKPALDKLGTGQAVSGPP-RKPPGTVA	523
Db	3265	EVGSPCRARRGGSTAHAPETDTADYIEPPK	-----SCVGCACSGPPEKKQOGAAE	331
Qy	524	GTARGPEGGTAQVAPAAASPPPEGVLTFTQSE	KMKMKELLVATKINSIAIKLQ	578
Db	3316	AEAPAPREGGAOEAPEGCTPEPAST----	GRTSKGKKKPLP-TKLPPSSVDLK	3365

[illegible]

Db 268 VKMGHAGSMGKVKVDNQHDQDIASV-----VALTKTVATAE---PFID-A 310
Qy 326 GYHCNRNRVREALQA-----WADTATVI----- 348
Db 311 KYDVRVQKIGNYKAYMRTSVSGNKNKTWTSAMLEQIAMSRYKLWVDTCSEIFGGLDIC 370
Qy 349 -----QDY-----NYCREDEEITYKEFFEIVANDVIPNL----- 375
Db 371 AVEALHGKDRHILIEVVGSSMPLIGDHODEKQLIVEL--VWNMAQALPRQRORDASP 428
Qy 376 -----LKPASALLBAG-----SQGTOSQGSALQD-----PE 408
Db 429 GRGSHGQTSPGALPLGRQTSQQPAGPPAQOQPPPPQGGPPQPGPQOQSGPLQOORPPQ 488
Qy 409 CFAHLLRYDGIKWEBSPTVLHVGNATFLVQSLGR----FEGQVQKRVIVSREAEA 464
Db 489 CQQLH-----SLGP-PAGSPLFQRLPSTSAQQPAPASQAAPTQGGGRQS-RPVAGGFGA 542
Qy 465 AEAEPEWGEAREGRRRGPP-----RRESKPEEPPPKPKALDKGLGTGQAVSGPPPKPPGT 521
Db 543 PPAARPPASPQO-RQAGPPQATQTSVSGPAPPKA-----SGAPPGGQOROGPPQKPPGP 597
Qy 522 VAGTARPEGGSTAQVPAPASPP-----PEGP 549
Db 598 AGPTRQASQAG-----PVPRTGPPPTQOQRPSPG 626

RESULT 11
D40750
proline-rich protein PRB1/25 (EA) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 03-May-1996
C:Accession: D40750
R:Azen, E.A.; Latreille, P.; Niece, R.L.
Am. J. Hum. Genet. 53, 264-278, 1993
A:Title: PRB1 gene variants coding for length and null polymorphisms among human saliva
A:Reference number: A40750; MUID:93304421
A:Accession: D40750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <AE>
A:Cross-references: GB:S62930
C:Superfamily: proline-rich protein

Query Match 3.7%; Score 118.5; DB 2; Length 117;
Best Local Similarity 40.0%; Pred. No. 0.22; Mismatches 28; Indels 13; Gaps 6;
Matches 34; Conservative 10; Mismatches 28; Indels 13; Gaps 6;

Qy 469 EPWGEAREGRRRGPRR--SKPEEPPPKPKALDKGLGTGQAVSGPPPKPPGTVAGTA 526
Db 12 KPOGPPPPPGKPGQPPPGGKNGKPGPPPPGKP--QG-PPQGDKSRSPSPPKPKPGPP 67
Qy 527 RGPEGGSTAQVPAPASPP--PEGP 549
Db 68 --POGGNPPQGP---PSPPGKPGQ 87

RESULT 12
I38346
elastic titin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: I38346
R:Labelit, S.; Koerner, B.
Science 270, 293-296, 1995
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330
A:Accession: I38346
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-7962 <RES>

A:Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g1017427
C:Genetics:
A:Gene: GDB:TTN
A:Cross-references: GDB:127867; OMIM:188840
A:Map position: 2q31-2q31

Query Match 3.7%; Score 118; DB 2; Length 7962;
Best Local Similarity 25.8%; Pred. No. 63;
Matches 48; Conservative 25; Mismatches 61; Indels 52; Gaps 8;

Qy 420 ICKWEESPTVLHVGNATFLVQSLGRFEGQV-RQKRVIVSREAEA---AEAEPEWGEAE 475
Db 6026 VAKKEAPPAKAVEPVQ-----KGVVTEEKITIVTQREESPPPAVPEIP-RKKV 6072
Qy 476 REGRRRPRRESKPEEPPPKPKALDKGLGTGQAVSGPPPKPPCTVACTARGEGGSTA 535
Db 6073 PEERKVPYRKEE--EVPKKPKVPAL-----PKKP-----PEEKVAV 6107
Qy 536 QVPAPASPPPEGVLTFQSEKMKMKELLVATKINSAIKLQLTAGSQVOMKKQKYSTP 595
Db 6108 PVPVAKKAPPAEVEV-----SKTVVEERKREVAEEKLSFAVQORVEVTRHVSAAE 6157
Qy 596 SDYTLS 601
Db 6158 EEWYSYS 6163

RESULT 13
PIHUPF
salivary proline-rich glycoprotein precursor PRB2 [validated] - human (fragment)
N:Alternate names: basic proline-rich peptide IB-8c precursor; proline-rich protein (N:Contains: basic proline-rich peptide IB-4; basic proline-rich peptide P-4
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1984 #sequence_revision 12-Apr-1996 #text_change 08-Dec-2000
C:Accession: E25372; A60827; A03294; B38355; A38355; F38355
R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A:Title: Differential RNA splicing and post-translational cleavages in the human saliv
A:Reference number: A92492; MUID:85289325
A:Accession: E25372
A:Molecule type: mRNA
A:Residues: 1-251 <MAE>
A:Cross-references: GB:K03208; NID:g190509; PIDN:AAA60189.1; PID:g190510
R:Manula, P.W.; Morley, D.J.; Larsen, S.H.; Karn, R.C.
Biochem. Genet. 26, 165-175, 1988
A:Title: Expression of human salivary protein genes.
A:Reference number: A60827; MUID:88240287
A:Accession: A60827
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 175-251 <MAM>
R:Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 93, 883-888, 1983
A:Title: Complete amino acid sequence of a basic proline-rich p-peptide, p-P, from huma
A:Reference number: A03294; MUID:83265674
A:Accession: A03294
A:Molecule type: protein
A:Residues: 134-194 <SAI>
R:Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
Biochemistry 30, 3351-3356, 1991
A:Title: Basic proline-rich proteins from human parotid saliva: relationships of the
A:Reference number: A38355; MUID:91190884
A:Accession: B38355
A:Molecule type: protein
A:Residues: 134-194 <KAU>
R:Experimental source: saliva
A:Experimental source: saliva
A:Note: this peptide, which is closely related to that of peptide p-E, contains three
A:Molecule type: protein
A:Residues: 10-67, R' <RA2>
A:Accession: F38355

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:17:57 ; Search time 21.22 Seconds
(without alignments)
1429.166 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFPLRSIDVV.....KVSTPDSYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 329605 seqs, 49716248 residues

Total number of hits satisfying chosen parameters: 329605

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Nev.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	5 US-09-948-941-267	Sequence 267, App
2	3185	100.0	610	5 US-09-380-337-2	Sequence 2, Appl
3	3185	100.0	639	5 US-09-948-941-387	Sequence 387, App
4	1604	50.4	341	5 US-09-758-466-660	Sequence 660, App
5	818	25.7	160	5 US-09-864-761-38608	Sequence 38608, A
6	147	4.6	994	1 PCT-US01-08631-31789	Sequence 31789, A
7	127.5	4.0	994	1 PCT-US01-08631-45699	Sequence 45699, A
8	123	3.9	331	5 US-09-538-092-845	Sequence 845, App
9	122	3.8	821	5 US-09-884-001-18	Sequence 18, Appl
10	118.5	3.7	1115	1 PCT-US01-14827-10666	Sequence 10666, A
11	118	3.7	705	5 US-09-317-063-2	Sequence 2, Appl
12	118	3.7	490	5 US-09-538-092-945	Sequence 945, App
13	118	3.7	711	1 PCT-US01-14826-313	Sequence 313, App
14	118	3.7	1273	1 PCT-US01-08631-43847	Sequence 43847, A
15	117.5	3.7	251	5 US-09-538-092-840	Sequence 840, App
16	115.5	3.6	140	5 US-09-834-366-13372	Sequence 13372, A
17	115.5	3.6	247	5 US-09-538-092-890	Sequence 890, App
18	115	3.6	227	1 PCT-US01-14827-8891	Sequence 8891, App
19	115	3.6	520	5 US-09-902-540-10545	Sequence 10545, A
20	114.5	3.6	131	5 US-09-834-366-13382	Sequence 13382, A
21	114.5	3.6	206	1 PCT-US01-14827-10568	Sequence 10568, A
22	111.5	3.5	1203	1 PCT-US01-08656-10232	Sequence 10232, A
23	111.5	3.5	1209	1 PCT-US01-08656-10231	Sequence 10231, A
24	111	3.5	276	5 US-09-538-092-889	Sequence 889, App
25	110.5	3.5	591	1 PCT-US01-08631-46126	Sequence 46126, App
26	110.5	3.5	591	1 PCT-US01-08631-48469	Sequence 48469, A
27	110.5	3.5	754	5 US-09-803-110-10319	Sequence 10319, A

28	110	3.5	300	1 PCT-US01-08631-50709	Sequence 50709, A
29	110	3.5	362	1 PCT-US01-14827-11292	Sequence 11292, A
30	110	3.5	511	1 PCT-US01-08631-49855	Sequence 49855, A
31	110	3.5	511	6 US-60-311-261-2486	Sequence 2486, App
32	110	3.5	807	1 PCT-US01-08631-45532	Sequence 45532, A
33	109.5	3.4	490	6 US-60-317-063-26	Sequence 26, Appl
34	109	3.4	755	1 PCT-US01-24104-57	Sequence 57, Appl
35	109	3.4	755	5 US-09-919-497-57	Sequence 57, Appl
36	109	3.4	1422	1 PCT-US01-08631-32704	Sequence 32704, A
37	108.5	3.4	481	5 US-09-948-941-353	Sequence 353, App
38	108.5	3.4	490	6 US-60-317-063-27	Sequence 27, Appl
39	108.5	3.4	864	5 US-09-538-092-1268	Sequence 1268, App
40	108.5	3.4	877	1 PCT-US01-14827-14448	Sequence 14448, A
41	108	3.4	179	5 US-09-758-440-788	Sequence 788, App
42	108	3.4	839	5 US-09-927-796-36	Sequence 36, Appl
43	108	3.4	1136	1 PCT-US01-14827-10458	Sequence 10458, A
44	108	3.4	1136	1 PCT-US01-08631-44700	Sequence 44700, A
45	108	3.4	1136	1 PCT-US01-08631-46239	Sequence 46239, A

ALIGNMENTS

RESULT 1

US-09-948-941-267
; Sequence 267, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-267

Query Match 100.0%; Score 3185; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGLKAAQKTLFPLRSIDVVRLFAAELGREEDLVLLSLVLGFEVHFPLAVNRVPTNPVE	60
DB	1	MGLKAAQKTLFPLRSIDVVRLFAAELGREEDLVLLSLVLGFEVHFPLAVNRVPTNPVE	60
QY	61	LTQPSPADPPGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGVSSRELVKK	120
DB	61	LTQPSPADPPGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGVSSRELVKK	120
QY	121	VSDVINLSLSRSYFKDRAHIQSLFSFTTGTCLDSSGVAFAVAVGACQALGLRDVHLALSED	180
DB	121	VSDVINLSLSRSYFKDRAHIQSLFSFTTGTCLDSSGVAFAVAVGACQALGLRDVHLALSED	180
QY	181	HAWVFGPNGEQTAAYTWHGKGNEDRRGQTVNAGVAERSWLYLKSGSYMRCDRKMEVAFMV	240
DB	181	HAWVFGPNGEQTAAYTWHGKGNEDRRGQTVNAGVAERSWLYLKSGSYMRCDRKMEVAFMV	240
QY	241	CATNPSTDLTDSLELLQLOQKLLWLLYDLGLHLETPMALGNLADLEELEPTGCRPDLT	300
DB	241	CATNPSTDLTDSLELLQLOQKLLWLLYDLGLHLETPMALGNLADLEELEPTGCRPDLT	300
QY	301	LYHKGTASAKTYRDEHIYPYMYLAGYHCRNRNVRALQAWADTATVIQDYNVCREDEEI	360
DB	301	LYHKGTASAKTYRDEHIYPYMYLAGYHCRNRNVRALQAWADTATVIQDYNVCREDEEI	360
QY	361	YKEFFEYANDVIPNLLKEAASLLEAGEERPGEQSQGTQSQSALQDPCEFAHLLRFYDGI	420

Db 361 YKEFFEYANDVIPNLLKEAASLLLEAGEERPEQSGTQSGSALQDPECFHLLRFYDGI 420
Qy 421 CKWEEGSPVPLVHVGWATFLVQSLGREGQVRQKRVIVSREAEAAAEPEWGEAREGRR 480
Db 421 CKWEEGSPVPLVHVGWATFLVQSLGREGQVRQKRVIVSREAEAAAEPEWGEAREGRR 480
Qy 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Db 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Qy 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVMKKQKVPSPDYTL 600
Db 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVMKKQKVPSPDYTL 600
Qy 601 SFLKQRKGL 610
Db 601 SFLKQRKGL 610

RESULT 2
US-09-380-337-2
; Sequence 2, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Settara C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Lubensky, Larisa V.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated with Multiple Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380.337
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-315100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-380-337-2
Query Match 100.0%; Score 3185; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLKAAQKTLFPLRSIDVVYRLFAAELGREPEDLVLLSLVGLGFVEHFLAVNRVPTNPVE 60
Db 1 MGLKAAQKTLFPLRSIDVVYRLFAAELGREPEDLVLLSLVGLGFVEHFLAVNRVPTNPVE 60
Qy 61 LTQPSAPAPPPGGLTYFPVADLSIIAALYARFTAQIRGAVDSLSPYREGVSSRELKVK 120
Db 61 LTQPSAPAPPPGGLTYFPVADLSIIAALYARFTAQIRGAVDSLSPYREGVSSRELKVK 120
Qy 121 VSDVINWNSLSRSYFKDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGRDVLHALSE 180
Db 121 VSDVINWNSLSRSYFKDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGRDVLHALSE 180
Qy 181 HAWVFGPNGEQTAEVTHWKGNDREDRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
Db 181 HAWVFGPNGEQTAEVTHWKGNDREDRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
Qy 241 CAINPSIDLHTDLSLELLQLOKLLWLLYDLGHLERYPMALGNLADLELEPTPGRPDPLT 300
Db 241 CAINPSIDLHTDLSLELLQLOKLLWLLYDLGHLERYPMALGNLADLELEPTPGRPDPLT 300
Qy 301 LYHKGASAKTYRDEHIYPYMLAGYHCRNRNVREALQAWADTATVIQDYNCREDEEI 360
Db 301 LYHKGASAKTYRDEHIYPYMLAGYHCRNRNVREALQAWADTATVIQDYNCREDEEI 360
Qy 361 YKEFFEYANDVIPNLLKEAASLLLEAGEERPEQSGTQSGSALQDPECFHLLRFYDGI 420
Db 361 YKEFFEYANDVIPNLLKEAASLLLEAGEERPEQSGTQSGSALQDPECFHLLRFYDGI 420
Qy 421 CKWEEGSPVPLVHVGWATFLVQSLGREGQVRQKRVIVSREAEAAAEPEWGEAREGRR 480
Db 421 CKWEEGSPVPLVHVGWATFLVQSLGREGQVRQKRVIVSREAEAAAEPEWGEAREGRR 480
Qy 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Db 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Qy 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVMKKQKVPSPDYTL 600
Db 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVMKKQKVPSPDYTL 600
Qy 601 SFLKQRKGL 610
Db 601 SFLKQRKGL 610

RESULT 3
US-09-948-941-387
; Sequence 387, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 387
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-387

Query Match 100.0%; Score 3185; DB 5; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.1e-185;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLKAAQKTLFPLRSDDVVRLFAAELGREGPDVLLSLVLGFVEHFLAVNRVPTNPVE 60
DB 30 MGLKAAQKTLFPLRSDDVVRLFAAELGREGPDVLLSLVLGFVEHFLAVNRVPTNPVE 89
QY 61 LTFQSPAPDPGGGLTYFPVADLSIIAALYARFTAGIRGAVDLSLYPRGGVSSRELVRK 120
DB 90 LTFQSPAPDPGGGLTYFPVADLSIIAALYARFTAGIRGAVDLSLYPRGGVSSRELVRK 149
QY 121 VSDVWNSLSRSYFKDRAHIQSLFSITGCKLDDSGVAFVAVGACQALGLRDVHLALSED 180
DB 150 VSDVWNSLSRSYFKDRAHIQSLFSITGCKLDDSGVAFVAVGACQALGLRDVHLALSED 209
QY 181 HAWVVGPNCEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFVY 240
DB 210 HAWVVGPNCEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFVY 269
QY 241 CAINPSIDLHTDSLELLQLOQKLLWLLYDLGHLERYPMALGNLADLELEPTPGRPDPLT 300
DB 270 CAINPSIDLHTDSLELLQLOQKLLWLLYDLGHLERYPMALGNLADLELEPTPGRPDPLT 329
QY 301 LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNRVREALQAWADTATVIQDNYCREDEEI 360
DB 330 LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNRVREALQAWADTATVIQDNYCREDEEI 389
QY 361 YKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPECFAHLLRYDGI 420
DB 390 YKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPECFAHLLRYDGI 449
QY 421 CKWEEGSPTPVLHVGNATFLVOSLGRFEGQVROKRVIVSREAEAAEAPWGEAEAREGR 480
DB 450 CKWEEGSPTPVLHVGNATFLVOSLGRFEGQVROKRVIVSREAEAAEAPWGEAEAREGR 509
QY 481 RGPRESKPEEPPPKKPKALDKGLTGQGVSGPPKPGTVAAGTARGPEGSTAOVPAP 540
DB 510 RGPRESKPEEPPPKKPKALDKGLTGQGVSGPPKPGTVAAGTARGPEGSTAOVPAP 569
QY 541 AASPPEGVPLTFQSEKMKMKELLVATKINSATKQLTAOSQVOMKKQKVTSPSDYTL 600
DB 570 AASPPEGVPLTFQSEKMKMKELLVATKINSATKQLTAOSQVOMKKQKVTSPSDYTL 629
QY 601 SFLKQKRGKL 610
DB 630 SFLKQKRGKL 639

RESULT 4
US-09-758-466-660
; Sequence 660, Application US/09758466
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM036
; CURRENT APPLICATION NUMBER: US/09/758,466
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 814
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 660
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-466-660

Query Match 50.4%; Score 1604; DB 5; Length 341;
Best Local Similarity 96.3%; Pred. No. 5.9e-90;
Matches 309; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

QY 291 PTPGRP-DPLTYLHKGIASAKTYRDEHIYPYMYLAGYHCRNRNRVREALQAWADTATVIQ 349
DB 24 PWPARTHELPLP---QCHCSAKTYRDEHIYPYMYLAGYHCRNRNRVREALQAWADTATVIQ 80
QY 350 DYNCREDEEIIYKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPEEC 409
DB 81 DYNCREDEEIIYKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPEEC 140
QY 410 FAHLRFYDGI CKWEEGSPTPVLHVGNATFLVOSLGRFEGQVROKRVIVSREAEAAEAE 469
DB 141 FAHLRFYDGI CKWEEGSPTPVLHVGNATFLVOSLGRFEGQVROKRVIVSREAEAAEAE 200
QY 470 PMGEAEAREGRRRGRPRRESKPEEPPPKKPKALDKGLTGQGVSGPPKPGTVAAGTARGP 529
DB 201 PMGEAEAREGRRRGRPRRESKPEEPPPKKPKALDKGLTGQGVSGPPKPGTVAAGTARGP 260
QY 530 EGGSTAQVAPAPASPPPEGPVLTFFQSEKMKMKELLVATKINSATKQLTAOSQVOMKK 589
DB 261 EGGSTAQVAPAPASPPPEGPVLTFFQSEKMKMKELLVATKINSATKQLTAOSQVOMKK 320
QY 590 QKVSTPDSYTL SFLKQKRGKL 610
DB 321 QKVSTPDSYTL SFLKQKRGKL 341

RESULT 5
US-09-864-761-38608
; Sequence 38608, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomico-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670

Qy	469	EPWGEAREGRRRCPRRE--SKBEP PPPKKPALDKGLGTGCGAVSGPRKkPGGTVAGTA	520
Db	46	KQCGPPPPKPGQGGPPQCGNKGPPGPKP---QC-PPPOGDKSRSPRSPKPGQGGP	101
Qy	527	RGPEGGSTAQVPAP-----AASPPEG	548
Db	102	--PQGGNQPGPPPPKPGQGGPPQ	125
RESULT 9			
US-09-884-001-18			
; Sequence 18, Application US/09884001			
; GENERAL INFORMATION:			
; APPLICANT: Bird, Timothy A.			
; APPLICANT: Peschon, Jacques J.			
; APPLICANT: Sims, John E.			
; APPLICANT: Virca, G. Duke			
; APPLICANT: Willis, Cynthia R.			
; TITLE OF INVENTION: Methods for Regulating Vascularization Using GEF			
; FILE REFERENCE: Containing NEK-Like Kinase (GNK)			
; CURRENT APPLICATION NUMBER: US/09/884,001			
; CURRENT FILING DATE: 2001-06-18			
; PRIOR APPLICATION NUMBER: 60/113,003			
; PRIOR FILING DATE: 1998-12-18			
; NUMBER OF SEQ ID NOS: 19			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 18			
; LENGTH: 821			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-884-001-18			
Query Match 3.88; Score 122; DB 5; Length 821;			
Best Local Similarity 20.2%; Pred. No. 11;			
Matches 98; Conservative 73; Mismatches 173; Indels 142; Gaps			
Qy	166	QALGLRDVHIALSDEHAWVFGPNGEOTAETVHWGKGNDRRGQTGVNAGVAERSWLYLKG	225
Db	254	QYISLNDNHISISVD--GLKFAEDGSE-----PNNDKMGHGHGPIVK-----LNG	298
Qy	226	SYMBCD-RKMEVAFMVCAINPSIDLHTD--SLELLQLOQKLLMLLYDLGLHLRYPMALGN	282
Db	299	DYRPTLRKGE-----SLNPVSDLFSELNISETKLQKQLM-----QVEREKAIL--	343
Qy	283	LADLEE-----LEPTPGRPDPLTYH-----KGIAKATYY-----RD--	315
Db	344	LANQESQTQLEHTKG---ALTEQHERVHRLTEHVNAMRGLOSKSELKAELEDGKCRDGS	400
Qy	316	EHYIPMY---LAGYHCRNR-NVREALQAWADTATVIQDNYNCRE---DEEI-YKEFFE	366
Db	401	EEADHYEVDINGLLEICKRYVAITEIDLKAEIKALKYKNKSVENYTDKAKYESKIQ	460
Qy	367	VANDVIPNLLKEAASLLEAGEERPEQ-----SQGTSGQSALQDPECFH-HLLR	415
Db	461	MYDEQVTSLEK---TTKESGKMAHMEKELQKMTSIANENHSTLNTAQDELVTFSLELAQ	517
Qy	416	FYDGICKWEBSPPVLHVGNATFVLSGRFEG-----QVRQKRVIVSREAEAAEA	467
Db	518	LYHWVLCNNETPNRVMLDYRQSRVTRFSGSLKGPDDPGRLLSPRLARRGVSSPVETR	577
Qy	468	EEPWGEAREGRRRCPRRESKPEPPPPKPALDKGLGTGCGAVSGPPRPGTVACTAR	527
Db	578	SEPAKASTE-----PSKEPSTKPTPTI-----	600
Qy	528	GPEGGSTAQVPAPASPPPEPGPPTFQSEKMGKMKELLVATKINSSAIKQLTAQSOVOM	587
Db	601	-----SPVITAPPSSPLVDTSDIRKEPNNIYNLNAIRIDQLKHLOKQAVDRSLQL	649
Qy	588	KKQKVS	593
Db	650	SRORAA	655

Db 311 KYDVRVOKIGONYKAYMRTSVSGNWKNTGSMLEQIAMSDRYKLWVDTCSIFGGLDIC 370
QY 349 -----QDY-----NYCREDEIYKEFFEYANDVIPNL----- 375
Db 371 AVEALHGKGRDHIIEVVGSSPLMGHDEKQDLVEL--VVNKAQALPRORQDASP 428
QY 376 -----LKEAASLEAG---EERPEQ-----SQGTQSGSALQD---PE 408
Db 429 GRGSHGQTSPGALPLGRQTSQAPGPAQORPPQGGPPQPGPQPGPQPLQQRPPQ 488
QY 409 CFAHLRLFDVGDCKWEGSPTVLHVGNATFLVQSIGR-----FEGQVROKVRIVSREAEA 464
Db 489 GOQHL-----SGLGP-PAGSPLPQLRPSPTSAPQQPASQAAPTQGGQGRS-REVAGGPGA 542
QY 465 AEAEPEWGEAREGRRGP---RRESKPEEPPPPKKPALDKGLGTGGGAVSGPPKPPGT 521
Db 543 PPAARPPASPSPO-ROAGPPQATRTQTSVSGPAPPKA-----SGAPPGGQQRQPPQRP 597
QY 522 VAGTARGPEGGSTAQPAPAAASP-----PEGP 549
Db 598 AGPTRQASQAG-----PVPRTGPTTQQPRPSGP 626

RESULT 13

PCT-US01-14826-313
; Sequence 313, Application PC/TUS0114826
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-103
; CURRENT APPLICATION NUMBER: PCT/US01/14826
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/577,298<151> 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781<151> 2000-10-24
; PRIOR APPLICATION NUMBER: 09/715,869<151> 2000-11-17
; PRIOR APPLICATION NUMBER: 09/775,330<151> 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 313
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-14826-313

Query Match 3.7%; Score 118; DB 1; Length 711;
Best Local Similarity 25.1%; Pred. No. 17;
Matches 52; Conservative 15; Mismatches 62; Indels 78; Gaps 9;
QY 411 AHLLR-FYDGICKWEGSPTVLHVGNATFLVQS-----LGRFEGQVROKVRIVSRE 461
Db 60 SHWARGFQNTCGPRGSPQPPRRPWASRVLQEATNWRAGPLAEVRAEQRKAASQE 119
QY 462 AEAAEAEPEWGEAREGRRGPRRESKPE-----EPPP 494
Db 120 REAKETERK--RRKAGGARRSPGPRPEPRNAPRVAQLAGLPAPLPERLAPVGRAPR 177
QY 495 PKPALDKG---LGTGGGAVSGGPR-----KPPGTVA----- 523
Db 178 SAQPSQSDPGSAWAGPWGRRGPRGPPSYEAHLLLRGSAGTAPRRWRDRPPPVAPPSEYEGPH 237
QY 524 ---GTARGPEGGSTAQP---APAAASP 544
Db 238 RTLGTRGP---GNSQVPTSSAPAAATP 261

RESULT 14

PCT-US01-08631-43847
; Sequence 43847, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 43847
; LENGTH: 1273
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (751)..(765)
; OTHER INFORMATION: Phage integrase family domain identified by eMATRIX,
; OTHER INFORMATION: accession number PF00589B, p-value=1.621e-11, raw score of 16.
; NAME/KEY: misc.feature
; LOCATION: (1)..(1273)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-43847

Query Match 3.7%; Score 118; DB 1; Length 1273;
Best Local Similarity 25.1%; Pred. No. 35;
Matches 52; Conservative 15; Mismatches 62; Indels 78; Gaps 9;
QY 411 AHLLR-FYDGICKWEGSPTVLHVGNATFLVQS-----LGRFEGQVROKVRIVSRE 461
Db 60 SHWARGFQNTCGPRGSPQPPRRPWASRVLQEATNWRAGPLAEVRAEQRKAASQE 119
QY 462 AEAAEAEPEWGEAREGRRGPRRESKPE-----EPPP 494
Db 120 REAKETERK--RRKAGGARRSPGPRPEPRNAPRVAQLAGLPAPLPERLAPVGRAPR 177
QY 495 PKPALDKG---LGTGGGAVSGGPR-----KPPGTVA----- 523
Db 178 SAQPSQSDPGSAWAGPWGRRGPRGPPSYEAHLLLRGSAGTAPRRWRDRPPPVAPPSEYEGPH 237
QY 524 ---GTARGPEGGSTAQP---APAAASP 544
Db 238 RTLGTRGP---GNSQVPTSSAPAAATP 261

RESULT 15

US-09-538-092-840
; Sequence 840, Application US/09538092
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 840
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P02812
US-09-538-092-840

Query Match 3.7%; Score 117.5; DB 5; Length 251;
Best Local Similarity 32.3%; Pred. No. 4.8;
Matches 31; Conservative 9; Mismatches 25; Indels 31; Gaps 5;
Qy 469 EPWGEAREGRRRGPRRE--SKPEEPKPPKALDKGLGTGGAVSGPP-----R 516
Db 25 QPQGPPPPGKPPQGPPPOGGNKPQGGPPPPGRP-----QGPPPOGDNKKSQAR 71
Qy 517 KPPGTVAGTARGEGGTAQVPAP----AASPPPEG 548
Db 72 SPPGKPPGPP--PQGGNPPGPPPPPPGKPPGPPPG 105

Search completed: October 12, 2001, 18:21:16
Job time: 199 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:17:22 ; Search time 90.47 Seconds
(without alignments)
1495.821 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFLRLSDDVV.....KVSTPDSYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters: 2803329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US060_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	12	US-08-865-337-6
2	2778	87.2	564	12	US-08-865-337-1
3	1124.5	35.3	813	23	US-60-171-625-525
4	1124.5	35.3	813	23	US-60-173-464-25061
5	1124.5	35.3	813	23	US-60-191-637-32161
6	1124.5	35.3	813	23	US-60-191-681-25486
7	1124.5	35.3	813	23	US-60-219-005-306
8	818	25.7	160	1	PCT-US01-00863-31783
9	619	19.4	120	23	US-60-196-710-5355
10	462.5	14.5	92	23	US-60-177-646-2142

11	342.5	10.8	72	23	US-60-177-646-2305	Sequence 2905, Ap
12	140.5	4.4	518	23	US-60-167-217-9939	Sequence 9939, Ap
13	140.5	4.4	518	23	US-60-173-464-8029	Sequence 8029, Ap
14	131	4.1	1151	14	US-09-023-905-4	Sequence 4, Appli
15	129	4.1	566	23	US-60-191-637-9944	Sequence 9944, Ap
16	129	4.1	566	23	US-60-191-681-7769	Sequence 7769, Ap
17	127	4.0	834	1	PCT-US01-08117-34	Sequence 34, Appli
18	127	4.0	834	23	US-60-188-986-34	Sequence 1481, Ap
19	125.5	3.9	603	23	US-60-230-435-1481	Sequence 101, App
20	123.5	3.9	720	1	PCT-US01-01567-101	Sequence 100, App
21	123.5	3.9	766	1	PCT-US01-01567-100	Sequence 61, Appli
22	123.5	3.9	973	1	PCT-US01-11988-1424	Sequence 1424, Ap
23	123.5	3.9	973	1	PCT-US01-11988-1424	Sequence 1424, Ap
24	123.5	3.9	973	22	US-09-833-245-1424	Sequence 21576, A
25	121	3.8	456	16	US-09-252-991A-21576	Sequence 1, Appli
26	121	3.8	930	23	US-60-280-260-1	Sequence 2473, Ap
27	119	3.7	705	18	US-09-488-725A-2473	Sequence 769, App
28	119	3.7	1291	23	US-60-242-679-769	Sequence 23911, A
29	118.5	3.7	501	18	US-09-417-507-23911	Sequence 6045, Ap
30	118.5	3.7	714	18	US-09-488-725A-6045	Sequence 41823, A
31	118	3.7	184	16	US-09-270-767-41823	Sequence 20, Appli
32	118	3.7	1336	1	PCT-US98-24950-20	Sequence 2, Appli
33	118	3.7	1336	13	US-08-975-424-20	Sequence 2, Appli
34	118	3.7	7962	17	US-09-312-216B-36	Sequence 2, Appli
35	117.5	3.7	124	1	PCT-US97-15799-2	Sequence 2, Appli
36	117.5	3.7	1581	15	US-09-110-517-2	Sequence 3, Appli
37	117.5	3.7	1581	15	US-09-110-517-3	Sequence 2, Appli
38	117.5	3.7	1581	17	US-09-346-231-2	Sequence 2, Appli
39	117.5	3.7	1581	17	US-09-346-231-3	Sequence 2, Appli
40	117.5	3.7	1581	17	US-09-346-231-2	Sequence 2, Appli
41	117	3.7	114	16	US-09-270-767-57068	Sequence 57068, A
42	117	3.7	114	16	US-09-270-849B-181000	Sequence 181000,
43	117	3.7	309	8	US-08-405-175-7	Sequence 7, Appli
44	117	3.7	696	15	US-09-129-668-4	Sequence 4, Appli
45	117	3.7	1442	23	US-60-229-518-325	Sequence 325, App

ALIGNMENTS

RESULT 1
US-08-865-337-6
; Sequence 6, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

```

, REFERENCE/DOCKET NUMBER: PF-0305 US
,
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415-855-0555
, TELEFAX: 415-845-4166
, TELEX:
,
, INFORMATION FOR SEQ ID NO: 6:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 610 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, IMMEDIATE SOURCE:
, LIBRARY: GenBank
, CLONE: 1945387
, US-08-865-337-6

```

```
Query Match          100.0%; Score 3185; DB 12; Length 610;
Best Local Similarity 100.0%; Pred. No. 2.8e-259;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	MGUKAAQKTLFPLP	URSIDDVV	RFPAEELGREEDP	VL	LSVLG	VFVEHFLAVNR	VITPNPE	60	
Db	1	MGUKAAQKTLFPLP	URSIDDVV	RFPAEELGREEDP	VL	LSVLG	VFVEHFLAVNR	VITPNPE	60	
Qy	61	LTQPSAPDP	PGGLTFY	PVADLS	II	TAALY	ARFTAQIR	CAVDLSLY	PREGGVSSREL	120
Db	61	LTQPSAPDP	PGGLTFY	PVADLS	II	TAALY	ARFTAQIR	CAVDLSLY	PREGGVSSREL	120
Qy	121	VSDVINWS	LSRSYFKDRAH	IQSLFS	ITGT	KL	DSSGAVAFV	YGACQALGLR	DVHLS	180
Db	121	VSDVINWS	LSRSYFKDRAH	IQSLFS	ITGT	KL	DSSGAVAFV	YGACQALGLR	DVHLS	180
Qy	181	HAWVFGP	NGEQTA	EVTHW	GKGNED	RRGQTV	NAGVA	ERSWLYL	KSGYMRCD	240
Db	181	HAWVFGP	NGEQTA	EVTHW	GKGNED	RRGQTV	NAGVA	ERSWLYL	KSGYMRCD	240
Qy	241	CAINPS	IDLH	TSLELL	QLOO	KL	LLWLYDL	GLH	ERYPMALGN	300
Db	241	CAINPS	IDLH	TSLELL	QLOO	KL	LLWLYDL	GLH	ERYPMALGN	300
Qy	301	LYHKGI	ASAKTY	YRDEH	ITPY	MYLAGY	HGCRNR	NVREALQ	AWDATV	360
Db	301	LYHKGI	ASAKTY	YRDEH	ITPY	MYLAGY	HGCRNR	NVREALQ	AWDATV	360
Qy	361	YKEFFE	VANDV	IENLL	KEAS	LL	EEGERP	EQSG	TSQGSALQ	420
Db	361	YKEFFE	VANDV	IENLL	KEAS	LL	EEGERP	EQSG	TSQGSALQ	420
Qy	421	CKWEGS	PTVL	HVGW	NATFLV	QSLGR	FEQGV	QVKRIV	SREAPAA	480
Db	421	CKWEGS	PTVL	HVGW	NATFLV	QSLGR	FEQGV	QVKRIV	SREAPAA	480
Qy	481	RGPR	RESKPE	PPPPK	PALD	KGLGT	GQGN	VSPP	PKPCTV	540
Db	481	RGPR	RESKPE	PPPPK	PALD	KGLGT	GQGN	VSPP	PKPCTV	540
Qy	541	AASPPP	EGPVL	TFQSE	KMGK	MELL	VATKIN	SSAI	LQLTA	600
Db	541	AASPPP	EGPVL	TFQSE	KMGK	MELL	VATKIN	SSAI	LQLTA	600
Qy	601	SFLK	RQRK	GL	610					
Db	601	SFLK	RQRK	GL	610					

RESULT

```

RESULT      2
US-08-865-337-1
; Sequence 1, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
;

```

```

Db 402 EVFANLRFYDGTGICWEDDSLTPIILHIGWAKPLVNNITRFDYDIRSQVVKLPEDLEAQ 461
Qy 458 -----VSREA-EAAAEAPGWEAREGRRR-----GPRRESKPEEPP-----493
Db 462 AKAEARAQEAKESKEAAGSEAMEGNNRMATKEERNVLALGKELSSLKPRNSQSSL 521
Qy 494 -----GTGQF-----GAVSGPP-----PPKKPALDKGL-----504
Db 522 LLINISSVLTTSTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSELTPTTLADLTAACG 581
Qy 505 -----GTGQF-----GAVSGPP-----RKPPGTV 522
Db 582 EKILNPDFLLGGGQCFADQKQPSGGESDNPELHNNNNNSNNNNNNHINADKKEAAT 641
Qy 523 AGTARGPEGGSTAQVPAPASPPG-----548
Db 642 TTNATTTSNGSTSVLPVSEANNAGQAQSQVQINDLGKPKQHKAKEETSDDDYDPFE 701
Qy 549 -----PVLTFQSEKMKMKELLVATKINSAILKQLTAQSOVMKKOKVSTPS 596
Db 702 IMLKRPVITLYSQMKMKGLDLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 4
US-60-173-464-25061
; Sequence 25061, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USBS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25061
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-25061

```

```

; ORGANISM: Drosophila
US-60-173-464-25061

Query Match          35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3 4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 1

Qy   9  TLFLRSIDDVVRILFAAEL-GREEPDVLVLSLVGLGVEHFLAVNRPVPTWNPBLTQPSP 67
      :||||| : ||| : |||| |::||::||| : ||| : ||| : ||| : ||| :
Db   10 SLFPLKSTADVINLFRALTSGIEPDLTLTSIVGYIELSLTTGAAQAQ-----AAQAAA 64

Qy   58 APDPGGITY-----FPVADLSITAAALYARFTAQIRGAVDLSLYPREGV 112
      | :||| : ||| : ||||| : ||| : ||| : ||| : ||| : ||| :
Db   65 AAVAAGDISQTATCGNDIIMGNVSPPVVTHELIAGLYKKRPQT-I LSV E---KPPIHQ 120

Qy   113 SSRELVKVSDVINWSLSRYSFKDRAHIQSLSFSITCTKLDDSSGFAPVAVGCAQLGRD 172
       :||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   121 ATREVIKKVSDVINWSLRYSKDRAHQLNLVSYLSGNKLDRCGVALAVVAGCOLLGKYKD 180

Qy   173 VHLALSEDHAWVFGPGNGEQTAETVTHWGKNEDRRGOTVNAGVAERSWLKYSYMKRCDR 232
      | :||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   181 VRLAISDEHWVFQCKRVETIEVTHWGKSEDKRGODIRPGIESGSWLVLGGVLAVVCER 240

Qy   233 KWEAFVAVCATNPSIDLHTDSLELLLOQKLLWLLYL DLGHLERYPMALGNLADLEELEPT 292
      ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   241 GMEVAACAAINISLTNSNSDCVAEVQQQLLWLLYDLGLHKLKRYPMALGT LGEELEIHR 300

Qy   293 PGRPDPLYHKGTASAKTYRDEHHYPMYLAGYHCNRNVREALQAWAL/TATVTDYN 352
      || : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db   301 HPSISCEOLYREATIESARTHYRNHHVVPYTYQGNYVNRLLKYRDFAFAWAANAADVIRLYT 360

Qy   353 Y-CREDDEIYKEFEVANVDVTPNLLKEAASLLEAGEBRRGEOSOGTOSOG----SALOQP 407
      | :||| : ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

Db 361 YQCRDDEIYKELLDIANELIPYVK-----TESSGHSARSLRDS 401
Qy 408 ECFALLRFYDGIKWEEGSPTPVLHVGMATFLVQSLGRPEGOVKVRI----- 457
Db 402 EVFANLLRFYDGIQWEEDSLTPILHIGWAKPLVNNITKFDYDIRSQVVIKLPEDLEAQ 461
Qy 458 -----VSREA-PAEAEPWGEAREGRR-----GPRRESKPEPP--- 493
Db 462 AKELARAQEAKEAKESKAAGSEAMEGNNRMATKEERNVNLGALGKELSSKPRNSOSSL 521
Qy 494 -----PPKKPALDKGL----- 504
Db 522 LLINISSVLTTETSYLLAVLTFLMLVHGLVSFAESAMVLYKSKNSELTPTTLADLTAACG 581
Qy 505 -----GTGQ-----GAVSGPP-----RKPGCTV 522
Db 582 EKILNPDLQGGGPPADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 641
Qy 523 AGTARGPEGGSTAOVAPAPASPPPEG----- 548
Db 642 TTNATTSGSGTSVOLPVSSEANNAQAOQSOVQINDOLGKPKQHEAKKEETSDDYDPFE 701
Qy 549 -----PVLTFQSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKOKVSTPS 596
Db 702 IMLKRPVITLYSQMKGLKDLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 5
US-60-191-637-32161
; Sequence 32161, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 32161
; LENGTH: 813
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-32161

Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

Qy 9 TFLPLRSIDDDVRLFAAEL-GREEPDVLVLSLVGFVHFVFLAVNRVPTNVPDLTFQSP 67
Db 10 SLFPLKSTADVINLFRALTSIGIEPDLTLLSIVVGYIELSLTTGEAAQ-----AAQAAA 64

Qy 68 APDPPGGITY-----FPVADLSIIAALYARETAQIRGAVDLSLYPREGV 112
Db 65 AAVAAGDISQATTGNDIIMGNSVPFPVTHELIAGLYKKFQT-ILSVVE---KPKPHRQ 120

Qy 113 SSRELKVKVDVINSLSRSYFKDRAHTQSLFSFTGPKLDSSGVAFVAVGACQALGLRD 172
Db 121 ATREVIKKVDVINSLSRSYKRAHLQNLVSYLSGNKLDKCFGVAVAVGCCOLLGYKD 180

Qy 173 VHLASEDHAWVFGPGEQTAETVTHGKGNDRRGQTVNAGVAERSWLYLKGSYMCRDR 232
Db 181 VRLAISEDHAWVFGQKRVETIEVTHGKGEDKRGQDIRPGIESGWSLYLGLLAVVCCR 240

Qy 233 KMEVAFMVCALNPISDLTHTDLSLELQLOKLLWLLYDLGLHLERYPMALGNLADLEELEPT 292
Db 241 GMEVAAICAAALNISLTSNSDCVAEALQQLLWLLYDLGLHLKRYPMALGTLGEEIHR 300

Qy 293 GPRDPLTYLHGKTASAKTYRDEHIYPYMYLAGYHCRNRVNRREALQAWDATATVIQYN 352
Db 300

Db 301 HPSISCEQLYREALIESARTHYRNHHVYPYTYQGYNYNRLKYRDAFAAWANNAADVIRLYT 360
Qy 353 Y-CREDEEIVYKERFEVANDVIPNLLKEAASLLEAGEERPGSQGTOSQ-----SAIQDP 407
Db 361 YQCRDDEIYKELLDIANELIPYVK-----TESSGHSARSLRDS 401
Qy 408 ECFALLRFYDGIKWEEGSPTPVLHVGMATFLVQSLGRPEGOVKVRI----- 457
Db 402 EVFANLLRFYDGIQWEEDSLTPILHIGWAKPLVNNITKFDYDIRSQVVIKLPEDLEAQ 461
Qy 458 -----VSREA-PAEAEPWGEAREGRR-----GPRRESKPEPP--- 493
Db 462 AKELARAQEAKEAKESKAAGSEAMEGNNRMATKEERNVNLGALGKELSSKPRNSOSSL 521
Qy 494 -----PPKKPALDKGL----- 504
Db 522 LLINISSVLTTETSYLLAVLTFLMLVHGLVSFAESAMVLYKSKNSELTPTTLADLTAACG 581
Qy 505 -----GTGQ-----GAVSGPP-----RKPGCTV 522
Db 582 EKILNPDLQGGGPPADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 641
Qy 523 AGTARGPEGGSTAOVAPAPASPPPEG----- 548
Db 642 TTNATTSGSGTSVOLPVSSEANNAQAOQSOVQINDOLGKPKQHEAKKEETSDDYDPFE 701
Qy 549 -----PVLTFQSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKOKVSTPS 596
Db 702 IMLKRPVITLYSQMKGLKDLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 6
US-60-191-681-25486
; Sequence 25486, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25486
; LENGTH: 813
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-25486

Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

Qy 9 TFLPLRSIDDDVRLFAAEL-GREEPDVLVLSLVGFVHFVFLAVNRVPTNVPDLTFQSP 67
Db 10 SLFPLKSTADVINLFRALTSIGIEPDLTLLSIVVGYIELSLTTGEAAQ-----AAQAAA 64

Qy 68 APDPPGGITY-----FPVADLSIIAALYARETAQIRGAVDLSLYPREGV 112
Db 65 AAVAAGDISQATTGNDIIMGNSVPFPVTHELIAGLYKKFQT-ILSVVE---KPKPHRQ 120

Qy 113 SSRELKVKVDVINSLSRSYFKDRAHTQSLFSFTGPKLDSSGVAFVAVGACQALGLRD 172
Db 121 ATREVIKKVDVINSLSRSYKRAHLQNLVSYLSGNKLDKCFGVAVAVGCCOLLGYKD 180

Qy 173 VHLASEDHAWVFGPGEQTAETVTHGKGNDRRGQTVNAGVAERSWLYLKGSYMCRDR 232
Db 181 VRLAISEDHAWVFGQKRVETIEVTHGKGEDKRGQDIRPGIESGWSLYLGLLAVVCCR 240

Qy 233 KMEVAFMVCALNPISDLTHTDLSLELQLOKLLWLLYDLGLHLERYPMALGNLADLEELEPT 292
Db 240

Db 241 GMEVAALCAALNLSLTSNSDCVEAEVLEQQQLLWLLVDLGHLLKRYPMALGTLGELIEIHR 300
QY 293 PGRPDPLTLVHKGIASAKTYRDEHIYPYMYLAGYHCRNVRNREALQAWADTATVIQDYN 352
Db 301 HPSISCQELYREALIESARTHYRNHHVPTYQGYNNRLLKYRDFAFAANAAADVIRLYT 360
QY 353 Y-CREDEEYKEFEFVANDVIPNLLKEAASLLEAGEERPGEGSQGTOSQ-----SALQDP 407
Db 361 YOCRDDEEYKELLDIANELIPYMK-----TESSGHSARSILRDS 401
QY 408 ECFALLRFYDGIKCKEESGPTVLHVGMATFVLSLGRFEGOVROKVRT----- 457
Db 402 EVFANLLRFYDGIKCKEESGPTVLHVGMATFVLSLGRFEGOVROKVRT----- 461
QY 458 -----VSREA-EAAAEAEPPGEEAREGRRR-----GPRRESKPEEP--- 493
Db 462 AKAEARAEQEAKEAESAAGSEAMEGNNRNMATKEERNVLAALGKLSKPRNSOSSL 521
QY 494 -----GTGQ-----GAVSGPP-----PPKKPALDKGL----- 504
Db 522 LLINISVLTETTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSELPPTTLADLTAACG 581
QY 505 -----GTGQ-----GAVSGPP-----RKPPGTV 522
Db 582 EKILNPDFLQGGQPFADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 641
QY 523 AGTARGEGGSTAQVPAPASPPPEG----- 548
Db 642 TTNATTTNSGTSVQLPVSEANNAGAQSOVQINDQLGKPOHKEAKKEETSDDYDPFE 701
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSOVQMKQKVSTPS 596
Db 702 IMLKRPVITLYSQMKGLKDLALLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 7

US-60-219-005-306
; Sequence 306, Application US/60219005
; GENERAL INFORMATION:
; APPLICANT: Skupski, Marion
; APPLICANT: et al.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS THAT SHOW
; TITLE OF INVENTION: HIGH HOMOLOGY TO HUMAN DISEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING THESE DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000734
; CURRENT APPLICATION NUMBER: US/60/219,005
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/171,625
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/175,870
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/187,207
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/191,639
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 306
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Drosophila
US-60-219-005-306

Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;
QY 9 TLFPRLSRIDVWRLFAAEL-GREEPDVLVLSVLGVFVEHFLAVNRVPTNPVLTQSP 67
Db 10 SLFPLKSTADVINFLRALTSIEPDLTLLSIVVGYIELTGTGAAQ-----AAQAAA 64

QY 68 APPPPGGGLTY-----FPVADLSIIAALYARFTAQIRGAVIILSYPPREGGV 112
Db 65 AAVAAGDISQATTCGGNDIIMGNSVPFVPTVTHLITAGLYKKFQT-ILSWE-----KPKPHRQ 120
QY 113 SSRELKVKSDVWLSLSRSFYFKDRAHIQSILFSFITCTKLDSSGVAFVAVGCAQALGRD 172
Db 121 ATREVTKKSDVWLSLSRSFYFKDRAHLQNLISYLSGNKLDGFCGVALAVVAGCOLLYK 180
QY 173 VHLALSDBHAWVYFGPNGEQTAETVTHGKGNEDRRGTVMAGVAERSWLVIKGSYMRCDR 232
Db 181 VRLAISDBHAWVYFGKRVETIEVTHGKGSDEKRGODIRPGIBSGSWLYLGLGLAVVCR 240
QY 233 KMEVAFMVCAINPSIDLHDSLELLQOQLLWLLYDLGHLERYPMALGNLADLEELEPT 292
Db 241 GMEVAALCAALNLSLTSNSDCVEAEVLEQQQLLWLLVDLGHLLKRYPMALGTLGELIEIHR 300
QY 293 PGRPDPLTLVHKGIASAKTYRDEHIYPYMYLAGYHCRNVRNREALQAWADTATVIQDYN 352
Db 301 HPSISCQELYREALIESARTHYRNHHVPTYQGYNNRLLKYRDFAFAANAAADVIRLYT 360
QY 353 Y-CREDEEYKEFEFVANDVIPNLLKEAASLLEAGEERPGEGSQGTOSQ-----SALQDP 407
Db 361 YOCRDDEEYKELLDIANELIPYMK-----TESSGHSARSILRDS 401
QY 408 ECFALLRFYDGIKCKEESGPTVLHVGMATFVLSLGRFEGOVROKVRT----- 457
Db 402 EVFANLLRFYDGIKCKEESGPTVLHVGMATFVLSLGRFEGOVROKVRT----- 461
QY 458 -----VSREA-EAAAEAEPPGEEAREGRRR-----GPRRESKPEEP--- 493
Db 462 AKAEARAEQEAKEAESAAGSEAMEGNNRNMATKEERNVLAALGKLSKPRNSOSSL 521
QY 494 -----GTGQ-----GAVSGPP-----PPKKPALDKGL----- 504
Db 522 LLINISVLTETTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSELPPTTLADLTAACG 581
QY 505 -----GTGQ-----GAVSGPP-----RKPPGTV 522
Db 582 EKILNPDFLQGGQPFADQKQPSGGSNDPELHNNNNNSNNNNNNHNAADKKEAAAT 641
QY 523 AGTARGEGGSTAQVPAPASPPPEG----- 548
Db 642 TTNATTTNSGTSVQLPVSEANNAGAQSOVQINDQLGKPOHKEAKKEETSDDYDPFE 701
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSOVQMKQKVSTPS 596
Db 702 IMLKRPVITLYSQMKGLKDLALLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 8

PCT-US01-00663-31783
; Sequence 31783, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)

QY 114 SRELVKVSDVIWNSLSRSYFKDRAHIQSLFSFITGTGKLDSSGVAF-----AVVGACQAL 168

; TELECOMMUNICATION INFORMATION:

C: 114 CDBY WJWZYCOCVETYNCT C OCUWZBBDCA * H : C Y C Q E F

QY 114 SRELVKVSDVIWNLSRSYFKDRAHIQSLFSFITGTKLDSSGVAF-----AVVGACQAL 168

Search completed: October 12, 2001, 18:20:49
Job time: 207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:14:02 ; Search time 13.8 seconds
(without alignments)
910.153 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFPLRSIDVV.....KVTSPSDYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	2	US-08-865-337A-6
2	2778	87.2	564	2	US-08-865-337A-1
3	117.5	3.7	124	2	US-08-925-237-2
4	117.5	3.7	1581	4	US-09-110-517-2
5	117	3.7	309	2	US-08-405-175A-7
6	117	3.7	696	3	US-08-906-865-4
7	114.5	3.6	535	4	US-09-029-348-1
8	114.5	3.6	537	4	US-09-029-348-4
9	108.5	3.4	2293	4	US-09-368-590-2
10	108	3.4	2205	1	US-08-093-453B-2
11	107.5	3.4	1321	2	US-08-317-310A-64
12	106.5	3.3	309	2	US-08-405-175A-8
13	106	3.3	928	5	PCT-US94-10357-3
14	105.5	3.3	82	2	US-08-925-237-4
15	105	3.3	501	2	US-08-969-630-4
16	105	3.3	816	1	US-08-038-760-3
17	105	3.3	816	2	US-08-470-091-3
18	105	3.3	928	1	US-08-204-329-1
19	105	3.3	928	2	US-08-959-638-8
20	105	3.3	928	2	US-08-482-627-5
21	105	3.3	928	3	US-08-801-092-4
22	105	3.3	928	4	US-08-328-673A-8
23	105	3.3	1009	2	US-08-357-642A-1
24	105	3.3	1009	2	US-08-460-626-1
25	104	3.3	502	2	US-08-969-630-2
26	104	3.3	534	4	US-09-029-348-5
27	104	3.3	661	2	US-08-795-868-14

28	104	3.3	928	5	PCT-US94-10357-2	Sequence 2, Appli
29	103	3.2	1149	3	US-08-560-005-5	Sequence 5, Appli
30	103	3.2	1479	3	US-08-840-062-2	Sequence 2, Appli
31	101.5	3.2	1418	3	US-08-963-825-20	Sequence 20, Appli
32	101.5	3.2	1418	4	US-09-010-999-1	Sequence 1, Appli
33	100.5	3.2	478	3	US-08-155-888-2	Sequence 2, Appli
34	100.5	3.2	1366	3	US-08-963-825-19	Sequence 19, Appli
35	100	3.1	1479	3	US-08-840-062-4	Sequence 4, Appli
36	99.5	3.1	565	4	US-08-961-083-218	Sequence 218, App
37	99.5	3.1	866	1	US-08-386-727-8	Sequence 8, Appli
38	99.5	3.1	866	2	US-08-600-452A-8	Sequence 8, Appli
39	99.5	3.1	1805	1	US-07-853-913-2	Sequence 2, Appli
40	99	3.1	1239	2	US-08-937-931-2	Sequence 2, Appli
41	99	3.1	1239	4	US-09-285-502-2	Sequence 2, Appli
42	98.5	3.1	258	4	US-08-961-083-90	Sequence 90, Appli
43	98.5	3.1	267	1	US-07-928-611-13	Sequence 13, Appli
44	98.5	3.1	267	2	US-08-487-811A-13	Sequence 13, Appli
45	98.5	3.1	267	5	PCT-US93-07370-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-865-337A-6
; Sequence 6, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1945387
US-08-865-337A-6

```
Query Match      100.0%; Score 3185; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-293;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDLVLSLVLFVFEHFLAVNRVPTNVPE 60
DB 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDLVLSLVLFVFEHFLAVNRVPTNVPE 60

QY 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
DB 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120

QY 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180
DB 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180

QY 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
DB 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240

QY 241 CAINPSIDLHTDSLLELLOQLKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300
DB 241 CAINPSIDLHTDSLLELLOQLKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300

QY 301 LYHKGIAASAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 360
DB 301 LYHKGIAASAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 360

QY 361 YKEFFEYVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 420
DB 361 YKEFFEYVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 420

QY 421 CKWEEGSPTPLVHVGWATFLVQSLGRPEGQVROKRVIVSREAEAAAEPEWGEAREGR 480
DB 421 CKWEEGSPTPLVHVGWATFLVQSLGRPEGQVROKRVIVSREAEAAAEPEWGEAREGR 480

QY 481 RGPRESKPEEPPPKPALDKGLGTGOGAVSGPPRPKPGTGTAGTARGPEGGSTAQVPAP 540
DB 481 RGPRESKPEEPPPKPALDKGLGTGOGAVSGPPRPKPGTGTAGTARGPEGGSTAQVPAP 540

QY 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKKQKVSTPSDYTL 600
DB 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKKQKVSTPSDYTL 600

QY 601 SFLKRQRKGL 610
DB 601 SFLKRQRKGL 610
```

```
RESULT 2
US-08-865-337A-1
; Sequence 1, Application US/08865337A
; Patent No. 5975649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murty, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337A
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14
CLONE: Consensus
US-08-865-337A-1

Query Match      87.2%; Score 2778; DB 2; Length 564;
Best Local Similarity 89.5%; Pred. No. 4.2e-255;
Matches 547; Conservative 2; Mismatches 14; Indels 48; Gaps 4;

QY 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDLVLSLVLFVFEHFLAVNRVPTNVPE 60
DB 1 MGLKAAQKTLFPLRSIDDDVRLFAAELGREPDLVLSLVLFVFEHFLAVNRVPTNVPE 60

QY 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
DB 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120

QY 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180
DB 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180

QY 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
DB 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240

QY 161 -----GPNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 214
DB 161 -----GPNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 214

QY 241 CAINPSIDLHTDSLLELLOQLKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300
DB 241 CAINPSIDLHTDSLLELLOQLKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300

QY 301 LYHKGIAASAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 359
DB 301 LYHKGIAASAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 359

QY 361 YKEFFEYVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 419
DB 361 YKEFFEYVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 419

QY 420 CKWEEGSPTPLVHVGWATFLVQSLGRPEGQVROKRVIVSREAEAAAEPEWGEAREGR 479
DB 420 CKWEEGSPTPLVHVGWATFLVQSLGRPEGQVROKRVIVSREAEAAAEPEWGEAREGR 479

QY 480 RGPRESKPEEPPPKPALDKGLGTGOGAVSGPPRPKPGTGTAGTARGPEGGSTAQVPAP 539
DB 480 RGPRESKPEEPPPKPALDKGLGTGOGAVSGPPRPKPGTGTAGTARGPEGGSTAQVPAP 539

QY 494 PAASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKKQKVSTPSDYT 553
DB 494 PAASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSOVOMKKQKVSTPSDYT 553

QY 600 LSFLKRQRKGL 610
```

```
Db 554 LSFLKRAQKGL 564
|||||
RESULT 3
US-08-925-237-2
; Sequence 2, Application US/08925237
; Patent No. 5981720
; GENERAL INFORMATION:
; APPLICANT: Azen, Edwin A.
; APPLICANT: Pan, David
; TITLE OF INVENTION: Human Salivary Proteins And Fragments
; TITLE OF INVENTION: Thereof Having Alpha-Glucosidase Inhibitory Activity
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kent Barta
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: USA
; ZIP: 53202-4108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,237
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Barta, Kent S.
; REGISTRATION NUMBER: 29,042
; REFERENCE/DOCKET NUMBER: 96429/9002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 283-2275
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-925-237-2

Query Match 3.7%; Score 117.5; DB 2; Length 124;
Best Local Similarity 32.3%; Pred. No. 0.0013;
Matches 31; Conservative 9; Mismatches 25; Indels 31; Gaps 5;

QY 469 EPWGEAREGRRRPRRE--SKPEEPPPPKPKALDKGLGTGOGAVSGPP-----R 516
Db 16 QPQGPVPPGKPGPPPPQGGNKKQGGPPPKP-----QGPPPGDNKSQAR 62
QY 517 KPQGTAGTARGEGGSTAQVPAP----AASPPPEG 548
Db 63 SPQKPGQGP--PQGGNQPGPPPPPKGKPGPPPG 96

RESULT 4
US-09-110-517-2
; Sequence 2, Application US/09110517A
; Patent No. 6248520
; GENERAL INFORMATION:
; APPLICANT: Roeder, Robert G
; APPLICANT: Fondell, Joseph D
; APPLICANT: Yuan, Chao X
; APPLICANT: Ito, Mitsuhiro
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NUCLEAR HORMONE
; TITLE OF INVENTION: RECEPTOR COACTIVATORS AND USES THEREOF
; FILE REFERENCE: 600-1-224
; CURRENT APPLICATION NUMBER: US/09/110,517A
; CURRENT FILING DATE: 1998-07-06
```

```
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-110-517-2

Query Match 3.7%; Score 117.5; DB 4; Length 1581;
Best Local Similarity 22.4%; Pred. No. 0.075;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDYNKYCREDEEYKFEFFVANDVIPNLLKEAASLLLEAGEERPEQSQ----- 395
Db 835 YTDPADLIADAAGSPSSDSPTNHHFFHDGVDFNPDL--NSQSQSGFGEEYFDESSQSGDND 893
QY 396 ----GTQSQ-----GSALQDPECFALHLLRF 416
Db 894 DFKGFASQALNTLGVPMLGDNGETKFKGNQADTVDFSIISVAGKALAPADLMEH---- 949
QY 417 YDGICKWEEGSPVPLVHVGWATFVQSLGRFEGOVROKVRIVSREAAEAEEFPGGEAR 476
Db 950 -----HSGSQGPLLTG-----DLGKEKTQKRYK---EGNGTNSNTLSGP-GLDSK 991
QY 477 EGRR-RGPRRESKPEEPPPKPALDKGLGTGOGAVSGPPKPPGTVAGTARGPEGGSTA 535
Db 992 PGKSRTPSNDGSKDKPKPKKADTEGKSPSHSS--SNRFTTP--TSTGSGSKSPGSAGRS 1049
QY 536 QVPAPAAAPP-----PEGPVLTTFQSEKMKGMKELLVATKINSSAIK-----LQL 579
Db 1050 QTPPGVATPPKTIQIPKGTVMV---GKPSSHSQYTSVSSGSSSGSKSHHSHSSSSSS 1106
QY 580 TAQSQVQMKKQKVSTPSTDYTLS 601
Db 1107 SASTSGMKMSSKSEGSSSSKLS 1128

RESULT 5
US-08-405-175A-7
; Sequence 7, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
```

TOPLOGY: linear

; Length: 553
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence

; TYPE: PRT
; ORCNTSM:

; ORGANISM: Artificial Sequence

; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: SEQUENCE
 ; OTHER INFORMATION: DERIVED FROM cDNA OF PROCOLLAGENS
 US-09-029-348-1

```

Query Match      3.6%; Score 114.5; DB 4; Length 535;
Best Local Similarity 34.2%; Pred. No. 0.026;
Matches 39; Conservative 5; Mismatches 31; Indels 39; Gaps

QY 473 EEAREGRR--RGPRRSKPEPP-----PPKKPAL-----DKGLG 505
      | : | | | | | | | | | | | | | | | | | | | |
Db 28 ETVRKPGADRGPRGERGPPGPRGDGDGPTGPPGPPGPPGLGNFAAQYDKGKG 87
      | | | | | | | | | | | | | | | | | | | |

QY 506 TGQGANVS-GPPRPDPGTVACTARGPEG-----GSTAQVPAPASPPGPGV 550
      | | | | | | | | | | | | | | | | | | | |

Db 88 LGFGPMGLMGPGRPPG--AAGAPGPGFGQFAGPEGPGGT- GPAGAPPHGPGV 138
      | | | | | | | | | | | | | | | | | | | |

```

```

RESULT      8
US-09-029-348--4
; Sequence 4, Application US/09029348
; Patent No. 6171827
; GENERAL INFORMATION:
; APPLICANT: THE VICTORIA UNIVERSITY OF MANCHESTER
; TITLE OF INVENTION: NOVEL PROCOLLAGENS
; FILE REFERENCE: G087857PUS LISTING
; CURRENT APPLICATION NUMBER: US/09/029,348
; CURRENT FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SEQUENCE
; OTHER INFORMATION: DERIVED FROM CDNA OF PROCOLLAGENS
US-09-029-348--4

```

```

Query Match      3.6%; Score 114.5; DB 4; Length 537;
Best Local Similarity 34.2%; Pred. No. 0.026;
Matches 39; Conservative 5; Mismatches 31; Indels 39; Gaps

Qy 473 EEARRR--RGPRRSKPEPP-----PPKKPAL-----DKGLG 505
      |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 28 ETVKGPAGDRGPRGERGPPGPRGDGEDGTPGPPGPPGLGGNFAAQYDGKGVG 87

Qy 506 TGQGVNS-GPGRPPTGTAGTARGPEG-----GSTAQVPAPASPPPEGV 550
      |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 88 LGPGPMGLMGPRGPPG--AAGAPGPGFGPAGPEGPGGT--GPACAPGPHGV 138
      |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

```

RESULT          9
US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:
; APPLICANT: Solimena, Michele
; TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
; TITLE OF INVENTION: ANTAGONISTS OF AUTOIMMUNE DISEASES
; FILE REFERENCE: 101918-200 (OCR-941)
; CURRENT APPLICATION NUMBER: US/09/368,590
; CURRENT FILING DATE: 1999-08-04
; EARLIER APPLICATION NUMBER: 60/095,657
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2293
; TYPE: PRT
; ORGANISM: Human

```

```

US-09-368-590-2

Query Match          3.4%; Score 108.5; DB 4; Length 2293;
Best Local Similarity 21.0%; Pred. No. 0.98;
Matches 92; Conservative 75; Mismatches 153; Indels 119; Gaps 25;

QY 168 LGLRDVHLALSEDH-----AWVFPGNGEQTAEVTHWKGKGNDRRGQTVNAGVA---ERS 219
Db 1336 LGEQEL-LMNSDEKGVKVRPSWCGCGPG-----GAGAGRRRLIVGALCPQDEQS 1384
QY 220 WLYLKGSYMRCDRKMEVAFWVCALNPSIDLHTDSLLELLOLQOQLLWLLDYDLGHLERYPMA 279
Db 1385 TLQLLKKHLOEQ-----GVENEESIAQLSRQCR--ALLENH----- 1421
QY 280 LGNLADLEELLEPTGPRDPLTLXHKGTASAKTYRDEHXYPMYLAGVHCNRRNVREALQ 339
Db 1422 ----PDEQISRRQSQVDRLYVALKEGEERRVALEQOQWLYQI-----SRVSE-LE 1469
QY 340 AWADTATVI-----QDYNCREDEEYTKEFFEVANDVIPNLLKEAASLLLEAGEERPGE 392
Db 1470 HWIAEKEVAGSPGLQDFEHVSLQEKSEF-----ASETGMAGRERLAA 1515
QY 393 QSQGTQSGSALQDP--ECFAHLRLFYDGICWKEGSGPTPVLHVGMATFLVQSLGRFEGQ 450
Db 1516 VNQ-----MVDELTEC-GHTAA--ATMAEWKDG-----LNEAWAE-LLELMG----- 1553
QY 451 VRQKRVIVSRAEA--AEAEFPWGEAEGRRRGRPRRESKPEEPPPPKKPALDKGLGTGQ 508
Db 1554 TRAQLLAASRLHKFFSDARELQ-GTEEKRRRLPRUTTPPE--PRPSASSMORTLRAFE 1610
QY 509 GAVSGPRKPGTVAGTARGEGGSTAQVPAPASPPPEGVLTFLQSEKMKMKELLVAT 568
Db 1611 HDLQ-----LLVSQVRQLQEGA-QRLRTVYAGEHAEA-IASREQEVLOCKWELLSAC 1660
QY 569 K-----INSSAIKQLLTAQ 582
Db 1661 EDARLHVSSADALRFHSQ 1679

RESULT 10
US-08-093-453B-2
; Sequence 2, Application US/08093453B
; Patent No. 5439814
; GENERAL INFORMATION:
; APPLICANT: Frey, Teryl K.
; APPLICANT: Dominguez, Geraldina
; APPLICANT: Wang, Chin_Yen
; TITLE OF INVENTION: Modified Infectious Rubella Virus
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jamie L. Greene, Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: United States
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50
; COMPUTER: Macintosh
; OPERATING SYSTEM: 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,453B
; FILING DATE: 19 JUL 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 07/722,334
; FILING DATE: 28 JUN 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 07362-0101

```


STATE: WI
COUNTRY: USA
ZIP: 53202-4108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,237
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Barta, Kent S.
REGISTRATION NUMBER: 29,042
REFERENCE/DOCKET NUMBER: 96429/9002
TELEPHONE: (608) 257-3501
TELEFAX: (608) 283-2275
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-925-237-4

Query Match 3.3%; Score 105.5; DB 2; Length 82;
Best Local Similarity 30.8%; Pred. No. 0.0092;
Matches 26; Conservative 10; Mismatches 18; Indels 31; Gaps 4;
QY 469 EPGEEAREGRRRGRPRE--SKPEPPPPKPKALDKGLGTGGGAVSGPPRRKPPGTVAGTA 526
DB 16 QPGPPPPKPKQPGPPPGGKNGKPGPPPGKP-----QGPP----- 51
QY 527 RGPEGGSTAQVPAPAA---SPPEEG 548
DB 52 --PQGNKSQGPppppgkpgpppg 74

RESULT 15
US-08-969-630-4
Sequence 4, Application US/08969630A
Patent No. 5981248
GENERAL INFORMATION:
APPLICANT: Xu, Hua
TITLE OF INVENTION: MAMMALIAN CELL DEATH PREVENTING KINASE, DPK
FILE REFERENCE: A463
CURRENT APPLICATION NUMBER: US/08/969,630A
CURRENT FILING DATE: 1997-11-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 501
TYPE: PRT
ORGANISM: mouse
US-08-969-630-4

Query Match 3.3%; Score 105; DB 2; Length 501;
Best Local Similarity 17.3%; Pred. No. 0.18;
Matches 92; Conservative 79; Mismatches 170; Indels 190; Gaps 23;
QY 151 KLDSSGVAFVAGACQ-----ALGLRDVHLALSDHAWVFGPNGEQTAEVTW 198
DB 53 QIEGTGIS---MSACKETALLRELNHPNVTALQKFLSHSDRKVWLLF---DYAKHDLW 105
QY 199 HGKGNDRRGQTWAGVAERSWLYLKGYSWRCRMEVAFMVCAINPSIDLHTDSLELLQ 258
DB 106 H-----IINFHASKA-----NKKPMQPKSMVKSLLYQLDGIHYLHAN----- 145
QY 259 LQQLKLWLY-----DLGHLERYPMALGNLADLELEPTGR 295

DB 146 -----WVLRDLKPANILVMGEGPERGRVKIADMGFARLNSPLKPLADLD----- 191
QY 296 PDPLTLYHKG---IASAKTYRDEHIYPMYLAG-----YHCRRNRNVK----- 335
DB 192 PVVTFWYRAPELLLGARHYTKAIDWAIGCIFAEILLTSEPIFHCROEDIKTSPFIHQD 251
QY 336 -----EALQAWADTATVIQDY---NYCREDEEYKKEFFEYANDVIFNLL 376
DB 252 LDRIFVMGFPADKDWEDIRKMPYPTLQKDFRRTTVANSSLIKYMEKUKVKPDSKVFL 311
QY 377 KEASLLEAGEERPEQSQGTQSGSALQDPECFALHLLRFYDGICKWEEGSPTVLIHWG 436
DB 312 LQKLLTMDPTKRITSEQ-----ALQDP-----YFOEDPLTLQV-- 345
QY 437 ATFLVQSLGRFEGQVROKRVIRVSREAEAEAEPEWGEAREGRRRRPRRESKPEEPPPK 496
DB 346 -----FAG-----CQIPYKREFLNEDEP--EE--KGDKNQFQQQNPHQQPAAPA 386
QY 497 KPALDKGLGTGGGAVSGPPRRKPP-----GTVAGTARGPEGST-----AQVPAPAA 543
DB 387 Q-----QTAAPPQAPPQSSAQTNGTAGGATAGGGGAGAGAGLQHSQDPGLNQV 434
QY 544 PPPEGPVLTFTQSEKMKMGKELLVATKINSSAIKQLTAQSQVQVMKKOKVST 594
DB 435 PPNKKPRIGPSGANSBG--PVMPSDYQHSSS---RLNYQSSVQSGSSQSQT 480

Search completed: October 12, 2001, 18:18:47
Job time: 285 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:13:42 ; Search time 22.09 Seconds
(without alignments)
1674.089 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFLPLRSIDVV.....KVSTPSTYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	19 AAW29749	Tumor suppressor p
2	2778	87.2	566	20 AAW86348	Human multiple end
3	131	4.1	1151	19 AAW77287	Zebrafish differ
4	126	4.0	580	22 AAB60493	Human cell cycle a
5	120.5	3.8	1318	22 AAB26874	HSV-2 immediate ea
6	118.5	3.7	1583	21 AAB29662	Human tyrosine pho
7	118.5	3.7	1636	21 AAB29661	Human histidine do
8	118	3.7	1301	20 AAW92296	Human alpha-1 (XVI
9	118	3.7	1336	20 AAY08694	Human collagen 18
10	117.5	3.7	124	19 AAW50192	Amino acid sequenc
11	117.5	3.7	1566	20 AAY06309	Human p53 regulato

12	117.5	3.7	1581	21 AAY69669	Human thyroid rece
13	117.5	3.7	1581	21 AAY69671	Human thyroid rece
14	117.5	3.7	1581	21 AAY69672	Human thyroid rece
15	116	3.6	2371	21 AAB41231	Human OREX ORF995
16	116	3.6	2972	22 AAB50363	Human SRCAP. Homo
17	116	3.6	3118	22 AAB50362	Human muscarinic a
18	114.5	3.6	479	10 AAW92406	Truncated pro-alpha
19	114.5	3.6	535	18 AAW12841	pro-alpha2(1):(111
20	114.5	3.6	537	18 AAW12840	Mammalian Ena (Men
21	114.5	3.6	541	19 AAW37148	Human collagen (Ty
22	113.5	3.6	1838	15 AAR53257	Human ALP. Homo s
23	113	3.5	1274	20 AAW89253	Human retinoblasto
24	111.5	3.5	928	22 AAB28270	HNRCR protein sequ
25	111.5	3.5	2453	21 AAB12454	Murine WASP protei
26	111	3.5	520	22 AAB49336	BetaIII spectrin p
27	111	3.5	2390	20 AAY05494	Breast cancer asso
28	110	3.5	754	20 AAY07027	Human OREX ORF1999
29	110	3.5	2220	21 AAY54004	Full length calcin
30	110	3.5	2220	21 AAY54004	Rat Neural-Wiskott
31	109.5	3.4	501	19 AAW46890	Human OREX ORF1813
32	109	3.4	551	21 AAB42049	Human EXMAD-7 SEQ
33	109	3.4	795	22 AAB27229	Human retinoblasto
34	109	3.4	928	20 AAY01543	Human activated T-
35	109	3.4	1851	22 AAB00023	Human OREX ORF1983
36	108.5	3.4	782	21 AAB42219	Amino acid sequenc
37	108.5	3.4	2293	21 AAY69197	Human SLAP-130. H
38	108	3.4	783	20 AAW70586	Retinoblastoma sus
39	108	3.4	928	18 AAW09411	Infectious rubella
40	108	3.4	2205	16 AAR79048	Mouse IRS-2. Mus
41	107.5	3.4	1321	17 AAR96994	Retinoblastoma ppl
42	107	3.4	928	16 AAR71680	Recombinant ppl10R
43	107	3.4	928	16 AAR71681	Plasmid pRF1 rat l
44	106.5	3.3	678	8 AAP70380	Human cytoskeleton
45	106	3.3	324	21 AAY91948	

ALIGNMENTS

RESULT 1
AAW29749
ID AAW29749 standard; Protein; 610 AA.
XX
AC AAW29749;
XX
DT 11-JAN-1999 (first entry)
XX
DE Tumor suppressor protein menin.
XX
KW Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;
KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO9839439-A1.
XX
PD 11-SEP-1998.
XX
PF 04-MAR-1998; 98WO-US04258.
XX
PR 05-MAR-1997; 97US-0040269.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
PI Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;
PI Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
XX
DR WPI; 1998-506360/43.
DR N-PSDB; AAV57415-16.
XX
PT Protein and gene associated with multiple endocrine neoplasia type 1

PT - useful in gene therapy and to diagnose sufferers of, and those
PT susceptible to, this condition by detecting protein absence or gene
PT mutation(s)

XX Claim 11; Page 56-57; 75pp; English.

CC This is the amino acid sequence of menin, a 67.5 kDa polypeptide
CC encoded by a novel human tumour suppressor gene, MEN1 (see AAV57416),
CC that is associated with multiple endocrine neoplasia type 1. Menin
CC exhibits no apparent amino acid sequence similarities to previously
CC known proteins. The lack of a functional menin polypeptide, either
CC by absence of the protein, its alteration and/or associated
CC mutations in the MEN1 gene, have been identified in individuals
CC suffering from multiple endocrine neoplasia type 1 (FEMEN1) and
CC identification of MEN1 provides a new window into the mechanism
CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
CC MEN1 associated cancers, and provides preclinical identification of
CC individuals with the FEMEN1 syndrome, i.e. cancer-free individuals
CC that are at high risk of acquiring MEN1 associated tumours. It
CC also provides a basis for gene therapy. Menin can be expressed in
CC eukaryotic or prokaryotic host cells and used to identify menin
CC specific reagents (e.g. antibodies) useful for detecting menin in
CC humans cells or tissues, especially from patients suspected of
CC being at risk from multiple endocrine neoplasia type 1.

XX Sequence 610 AA;

Query Match 100.0%; Score 3185; DB 19; Length 610;
Best Local Similarity 100.0%; Pred. No. 3.5e-284;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLKAAQKTLPLRSIDVYVRLFAELGREGPDVLVLSLVGFVEHFLAVNRVPTNPVE 60
Db 1 mglkaaqktlplsidsdvrlfaaelgreepdlvllslvlgfvehflavnrviptnpve 60
Qy 61 LTFQSPAPPPGGTLTFPPVADLSIIAALYARFTAQIRGAVDSLXPRGGVSSRELVKK 120
Db 61 ltfqspapppggltfppvadlsiaalyarftaqirgavdlslypregvssrelvkk 120
Qy 121 VSDVWNSLSRSYFKRAHQTSLSFTGKFLDSSGVAFVAVGACQALGRDVLHALSED 180
Db 121 vsdvwnslsrsyfkdrahqtslsftgkfldssgvafavvgacqalgrdvhlalse 180
Qy 181 HAWVVFPGNGSOTAEVTHWKGNDRRGQTVNAGVAERSWLYLKGYMRCDRKMEVAFMV 240
Db 181 hawvvfpgngeqtavtwhkgndrrgqtvnagvaerswlylkgymrcdrkmevafmv 240
Qy 241 CAINPSIDLHFTDSLLELQLOKLLWLWLYDLGLHLEPYPMALGNLADLELEPTGRPDPLT 300
Db 241 calnpsidlhtdslellqqlkllwlydlghlerypmaignladleleptgrpdplt 300
Qy 301 LYHKGIASAKTYRDEHYIPYMYLAGYHCRNRVREALQAWADTATVIQDYNCREDEEI 360
Db 301 lyhkgiasaktyrdehyipymylagyhcnrnrvrealqawadtatviqdyncredee 360
Qy 361 YKEFEVANDVIPNLLKEASLLLEAGEERPGESQGSQSGSALQDPCEFAHLLRFYDGI 420
Db 361 ykefevandvipnllkeasllleageerpgesqgsqgsalqdpcecfahllrfydg 420
Qy 421 CKWEGSPTVLYHWGATFLVQSLGREGQVRQKRVIRVSREAAEAEPWGEAREGR 480
Db 421 ckweegsptvlyhwgatlflvqslgreqvqrkvrirvsreaaeaepwgearegr 480
Qy 481 RGPRESKPEPPPPKPKALDKGLTGCGAVSGPPRPGTIVAGTARGCGSTAQVPAP 540
Db 481 rgpreskpeppppkpkaldkgltgcgavsgpprpgtvtvagtargcgstaqvpap 540
Qy 541 AASPPPEPVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSOVMKQKVSFSDYTL 600
Db 541 aaspppepvltfqsekmgkmkellvatkinsiaiklqtaqsovmkqkvsfspdtytl 600

Qy 601 SFLKRRQRKGL 610
Db 601 sflkrqrkgl 610

RESULT 2
AAW86348
ID AAW86348 standard; Protein; 566 AA.
XX AC AAW86348;
XX 15-MAR-1999 (first entry)
XX Human multiple endocrine receptor neoplasia type 1 protein.
XX Human; multiple endocrine neoplasia type 1 protein; MEN1-1;
XX tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
XX detection; diagnosis; drug screening.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Misc-difference 137 /label= unknown
XX Misc-difference 161 /note= "encoded by YGG"
XX Misc-difference 274 /label= unknown
XX Misc-difference 274 /note= "encoded by NNT"
XX Misc-difference 312 /label= unknown
XX Misc-difference 417 /note= "encoded by SCC"
XX Misc-difference 421 /label= unknown
XX Misc-difference 424 /note= "encoded by NGC"
XX Misc-difference 428 /label= unknown
XX Misc-difference 435 /note= "encoded by CNG"
XX Misc-difference 443 /label= unknown
XX Misc-difference 443 /note= "encoded by NCC"
XX WO9854324-A1.
XX 03-DEC-1998.
XX 29-MAY-1998; 98WO-US10957.
XX 29-MAY-1997; 97US-0865337.
XX (INCY-) INCYTE PHARM INC.
XX Au-Young J, Covitz PA, Murry LE, Tang YT;
XX WPI; 1999-059839/05.
XX N-PSDB; AAV80659.
XX New isolated multiple endocrine neoplasia type 1 protein - used to
XX develop products for the diagnosis, treatment and prevention of
XX cancers and multiple endocrine neoplasia.
XX Claim 1; Fig 1; 67pp; English.

XX The present sequence is human multiple endocrine neoplasia type 1
 CC protein (MEND-1). MEND-1 plays a role in multiple endocrine neoplasia
 CC when one or both normal genetic copies of MEND-1 are mutated and no
 CC longer able to suppress tumorigenesis. MEND-1 can be used for treating
 CC or preventing cancers and multiple endocrine neoplasia. MEND-1
 CC polynucleotides can also be used for gene therapy. Products from the
 CC present invention can also be used for detection, diagnosis and drug
 CC screening.
 XX
 SQ Sequence 566 AA;

Query Match 87.2%; Score 2778; DB 20; Length 566;
 Best Local Similarity 89.5%; Pred. No. 8.7e-247;
 Matches 547; Conservative 2; Mismatches 16; Indels 46; Gaps 4;

QY 1 MGLKAQKTLFPLRSIDDDVRLFAAELGREEDLVLLSLVLGFEVHFLAVNRVIPNVPE 60
 DB 1 mglkaeqktlflrsiddvrlfaaelgreepdlvllslvlgfvehflavnrviptnvpe 60
 QY 61 LTFQSPADPPGGLYFYPVADLSIAALYARETAQIRGAVDSLPLPREGVSSRELKVK 120
 DB 61 ltfqspadpppgglyfypvadislaalyaretaqirgavdlslypreggvssrelvkk 120
 QY 121 VSDVINSLRSYFKDRAHQSLFSFTITGKLDSSGVAFVAVGACQALGLRDVHLASED 180
 DB 121 vsdvinslrsyfkdxahqslfsftitgkld-sgvafa----- 159
 QY 181 HAWVFGNGEQAEVTHGKGNEDERRGQTVNAGVAERSWLYLKSGYMRCDRMEVAFMV 240
 DB 160 -----vxgpngeqtaevthgkgnedrrgqtvnagvaerswlylksgymrcdrkmevafmv 215
 QY 241 CATNPSTDLTDSLELLOQKLLWLLYDLGHLERYPMALGNLADLEELEPTGCRDPLT 300
 DB 216 catnpsldltdslellqgkllwlydlghlerypmaignladleeleptgprpoxa 275
 QY 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVREALQAWA-DTATVIQDYNCREDEE 359
 DB 276 llhkgiasaktyrdehiypymylagyhcrnrnvrealqalgrdtatliqdyncredee 335
 QY 360 IYKEFEVANDVPNLLKEAASLLERGERPGESQSGTQSQGALQDPECFALLRFPYDG 419
 DB 336 iykefevandvipnlkkaaslleageerpgesqsgtqsgsalqgpecfahllrfydg 395
 QY 420 ICKWEGSPVPLVHGWAFLVQSLGRFEGQVQKRVIRVSREAEAEAEPEWGEAREGR 479
 DB 396 ickweegspcvplhvgwaflixgslrxfqgqvkvrvirvrdaea----- 441
 QY 480 RRGPRRESKPEEPPPKPALDKGLGTGQGVAGSGPPRKPPTVAGTARGPESGTAQVPA 539
 DB 442 -----exkpeepppkpaldkglgtgqgvagsgpprkpvtvagtargpegsgtaqva 495
 QY 540 PAASPPPEGVPLTFQSEKMGKMKELLVATKINSALKQLTAQSQVOMKKQKYSTPDSYT 599
 DB 496 paaspppegvpltfqsekmgmekellvatkinsalkqltaqsqvqmkqkvkstpsdyt 555
 QY 600 LSLFKRQKGL 610
 DB 556 lsflkrqakgl 566

RESULT 3
 AAW77287

ID AAW77287 standard; Protein; 1151 AA.

XX

AC AAW77287;

XX 20-NOV-1998 (first entry)

DE Zebrafish differentiation enhancing factor 1 protein.

XX Zebrafish; differentiation enhancing factor; ankyrin repeat; C2 domain;
 KW

KW SH3 consensus binding sequence; pleckstrin homology domain; adipogenesis;
 KW neurogenesis; hyperplastic disease; neoplastic disease; nervous system.
 XX Brachydanio rerio.
 OS
 XX
 PN W09836065-A1.
 XX
 PD 20-AUG-1998.
 XX
 PF 13-FEB-1998; 98WO-US02724.
 XX
 PR 14-FEB-1997; 97US-0038191.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Chan J, Harris DF, Hu E, King FJ, Spiegelman B;
 PI Thomas RM;
 XX
 DR WPI; 1998-467173/40.
 DR N-PSDB; AAV59104.
 XX
 XX New nucleic acid encoding differentiation enhancing factor - used
 PT particularly to regulate adipogenesis and neurogenesis, e.g. for
 PT treating tumours and neurological disease
 XX
 PS Claim 3; Fig 12; 203pp; English.

XX The differentiation enhancing factors (DEF), comprise at least one each
 CC of SH3 consensus binding sequence, ankyrin repeat, pleckstrin homology
 CC domain and C2 domain. The DEF induce adipogenesis or neurogenesis; they
 CC are mediators of SH3-domain dependent signalling and may be involved in
 CC cellular gene expression, cytoskeletal architecture, protein trafficking,
 CC endocytosis or adhesion, migration, proliferation and differentiation of
 CC cells. Typical applications of DEF and agents that modulate
 CC interaction between the protein and it's ligand, or of nucleic acid
 CC expressing them, are treatment of hyperplastic and neoplastic diseases
 CC (a wide range of solid tumours and leukaemias), including metastases; for
 CC in vitro induction of differentiation of neural crest cells to neurons,
 CC glial cells etc.; for increasing neuron survival, and inducing cell
 CC repair, in the nervous system (e.g. treatment of traumatic injury,
 CC stroke, Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic
 CC lateral sclerosis, multiple sclerosis etc.).

Sequence 1151 AA;

Query Match 4.1%; Score 131; DB 19; Length 1151;
 Best Local Similarity 21.4%; Pred. No. 0.0097;
 Matches 118; Conservative 73; Mismatches 220; Indels 140; Gaps 26;

QY 119 KKYSDVTWNSLSRSYF---KDRAHIQSLFSFTITGKLDSSGVAFVAVGACQALGLRDVHL 175
 DB 382 kkcfdli--shnrythfqaedegefviwsvltnskealnmafrgeqsagdsldltk 439
 QY 176 ALSDEHAWVFGPNGEQAEVTHGKGNEDERRGQTVNAGVAERSWLYLKSGYMRCDRME 235
 DB 440 aliiedvlri-----pgne-----vccdcgvpepkwlstnlgilc----- 474
 QY 236 VAFWVCA-INPSTDLH---TDSLELLOQKLLWLLYDLGHLERYPMALGNLADLEELEP 291
 DB 475 ---iecsghremghisriqsmeldkigtseillaknvgvsnssfnleegn-----p 524
 QY 292 TPG-RPDPDLTYHKGIASAKTYRDE---EHIYPYMYLAGYHCRNRNVREALQAWADTATV 347
 DB 525 spspkpa-----ssdmterkeyinakyvehrfarrtattatargdlyeavrt-rdlmal 579
 QY 348 IQDYNCREDEEITYKEFEVANDVIPNLL-----KEAASLLEAGEERKPEGSQGTQS 399
 DB 580 iqly---adgvelmdpfpeagqdpgetalhfvartsqtslhldvflvqnsqgtdrqt 636
 QY 400 QGSAL-----QDPECFAHLLRFYDGICKWEGSPTP-----VLHVGWATFLVQ-SLGR 446
 DB 637 gnaalhyccyckpecikllilrgkpsidlvnqngetaldiarlrnvqceallveaagr 696

XX PS Claim 1; Page 17-21; 28pp; English.

XX CC This invention relates to an immediate early herpes-simplex-virus type 2

CC (HSV-2) infected cell protein 4 (ICP4) recognised by human cytotoxic T

CC cells. HSV-2 ICP4 protein is recognized by cytotoxic T-lymphocyte (CTL)

CC cells in humans and is used in vaccines for therapeutically or

CC prophylactically treating HSV infections. Pharmaceutical compositions of

CC HSV-2 ICP4 protein may be used to treat patients suffering from HSV

CC infections, to prevent or decrease recurrent herpes disease, frequency,

CC severity and duration of episodes. The present sequence represents the

XX HSV-2 ICP4 amino acid sequence of the invention.

SQ Sequence 1318 AA;

Query Match 3.8%; Score 120.5; DB 22; Length 1318;

Best Local Similarity 25.4%; Pred. No. 0.11;

Matches 62; Conservative 19; Mismatches 108; Indels 55; Gaps 9;

QY 354 CREDEIYKEFEVANDVIPNLLKEAASLLEAGEERPGESQGTOSQSALQDPECFAHL 413

DB 644 crgilealaegfdgdlaavpgl----agarpaaprrpgagaaphada---prlrawl 596

QY 414 --LRF-----YDICKWEEGPTPLVHVGMATFLVOSLG----- 445

DB 697 relfrvdalvmlrldrvaggseaaavaavrsavslvagalgpalprsprllssaaaaa 756

QY 446 ---RFGQVRQV---RIVSRGAEEAEEPEWGEAREGRRGPRRESKPEPPPKK-- 497

DB 757 adllfqnglrplladtvaadslapasap--rearkrksaparapggagprppkkr 814

QY 498 ----PALDKGLGTGGGAVSGPPRPPGVAGTAR-----GPEGGTAQVPAPAAASPPPE 547

DB 815 adaprpaaappagaappaptpprprpaaltrtpaegpbgggwrrpgpgshtpaps 874

QY 548 GPVL 551

DB 875 aaal 878

RESULT 6

ID AAB29662 standard; Protein; 1583 AA.

XX AC AAB29662;

XX DT 23-FEB-2001 (first entry)

XX DE Human tyrosine phosphatase HD-PTP fragment.

XX KW Human; histidine domain-protein tyrosine phosphatase; HD-PTP;

XX KW chromosome 3p21.3; gene deletion; tumour suppressor; cytostatic;

XX KW lung cancer; tumour; gene therapy; diagnosis; recombinant production;

XX KW anticancer.

XX OS Homo sapiens.

XX PN WO200063392-A1.

XX PD 26-OCT-2000.

XX PF 14-APR-2000; 2000WO-JP02455.

XX PR 16-APR-1999; 99JP-0108842.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Shimizu K;

XX WPI; 2000-672740/65.

XX N-PSDB; AAC81225.

PT Human tyrosine phosphatase with oncostatic activity encoded by a gene

PT frequently deleted in lung cancer, useful for treatment and diagnosis

PT of tumors

XX Example 2; Page 82-98; 134pp; Japanese.

XX The invention relates to a novel human tyrosine phosphatase, histidine

CC domain-protein tyrosine phosphatase (HD-PTP; AAB29661) and to human

CC HD-PTP nucleic acids (AAC81224, AAC81225, AAC81262, AAC81263). The

CC HD-PTP gene is located on chromosome 3p21.3. This region is frequently

CC found to be deleted in lung cancers, and is therefore thought to contain

CC a tumour suppressor gene. The invention also relates to expression

CC vectors and host cells containing human HD-PTP nucleic acids; the

CC recombinant production of HD-PTP; anticancer drugs containing HD-PTP;

CC gene therapy compositions containing DNA encoding HD-PTP; diagnostic

CC reagents containing HD-PTP oligonucleotides; antibodies specific for

CC HD-PTP; and an immunoassay method using HD-PTP-specific antibodies for

CC use in cancer diagnosis and investigation. HD-PTP proteins, nucleic acids

CC and antibodies may be used in the treatment, investigation and diagnosis

CC of cancers, particularly those of the lung. The present sequence

XX represents a substantial portion of the human HD-PTP protein.

SQ Sequence 1583 AA;

Query Match 3.7%; Score 118.5; DB 21; Length 1583;

Best Local Similarity 21.8%; Pred. No. 0.22;

Matches 138; Conservative 73; Mismatches 224; Indels 197; Gaps 34;

QY 24 AAELGREPDVLLSL-----VLG-----FVEHFLAVNRVPTNVPETLTFQP- 65

DB 230 aiklakgqdtvgdalrftmdvlggkynsakkndfiyh-----eavpalclqpv 280

QY 66 --SPAPDP--PGGLTFPPVADLSIIAALYARFQAIRGAVDSLTPREGGVSSRELKVKYS 122

DB 281 kgaplvpkplvnpdpavtvgpddifaklvpmaheas-----slyseekaklremmakie 335

QY 123 DVIWNSLSRSYFKDRAHIQSLFSFITGTIKLDSGV-----AFA-----VVGACAL----- 168

DB 336 d-----knevlqdfmsmqldpetvndldayshippqlmekcaalsvrpd 380

QY 169 -----GLRDVHLALSEDHAWVFGPNCEQTAETVTHGKGNEDDR 207

DB 381 tvrnlvqsmqvlgvftdveaslkdirleed-----elleqkfgevvgqagai 430

QY 208 GQTVNAGVAE--RSWLYLKGSYMRCDKMEVAFWCAINPSIDLHTDLSLELLOLQKLLW 265

DB 431 sitskaelaevrrew----akymevhek--asftnselhramnlhvgnlrlls----- 477

QY 266 LLYDLGHLERYPMALGNLADLELEPTP--GRDPDLTLVH-KGIASAKTYVYRDEHIVPYM 322

DB 478 -----gpldqvraal-----ptpalspedkavqlgnlrlakvqemtdqrv---- 518

QY 323 YLAGYHCRNRNVREALQAWADATVI-QDYNVYCRE-DEEIIYKEFEFVANDVIPNLLKE-- 378

DB 519 -----slegqlreliqkdditaslvttdhsemkklifeeqkkydqklyleqnlqaadr 572

QY 379 -AASLLEAGEERPEEQ---SOGTOSGSALQ---DPECFAHLRLRYDGTCKWEESGPTP 430

DB 573 vicalteanvgaavrvrvisldldqkwnstlqtlvasyayedlmk-----ksqegrdfy 626

QY 431 V-LHVGWATFLVQSLGRFEGQVRQKVRIVSR-----EAEAAEAE 468

DB 627 adleskvaallertgstcqareaaqqldrelkkpprptapkpplprreseaveag 686

QY 469 EPWGEARE---GRRRGPR-----RESKPEPPPPKPKPALDKGLGTGGGAVSGPPRKP 519

DB 687 dp-peeelrsllpdmvagrplpdtflgsatplhfpssfpfs---stgpgplylsqp--lpp 740

QY 520 GIVAGTARGPEGGSTAQVPAPAAASPPPEGPVL 551

DB 741 gtyS-----gptqliqpragpphampvapgal 768

```

RESULT      7
AAB29661
ID AAB29661 standard; Protein; 1636 AA.
XX AC
AAC BAA08516
XX AC AAB29661;
XX DT 23-FEB-2001 (first entry)
XX XX Human histidine domain-protein tyrosine phosphatase, SEQ ID NO:2.
DE DE
XX XX Human; histidine domain-protein tyrosine phosphatase; HD-PTP;
KW chromosome 3p21.3; gene deletion; tumour suppressor; cytostatic;
KW lung cancer; tumour; gene therapy; diagnosis; recombinant production;
KW anticancer.
XX OS Homo sapiens.
XX PN WOI000063392-AI.
XX PD 26-OCT-2000.
XX PF 14-APR-2000; 200OWO-JP02455.
XX PR 16-APR-1999; 99JP-0108842.
XX PP (KYOW ) KYOWA HAKKO KOGYO KK.
XX PI Shimizu K;
XX DR WIPI : 2000-672740/65.
XX DX N-PSDB ; AAC81224 .
XX PT
XX PT Human tyrosine phosphatase with oncstatic activity encoded by a gene frequently deleted in lung cancer, useful for treatment and diagnosis of tumors -
XX PS Claim 1; Page 69-77; 134pp; Japanese.
XX CC The invention relates to a novel human tyrosine phosphatase, histidine domain-protein tyrosine phosphatase (HD-PTP; AAB29661) and to human HD-PTP nucleic acids (AAC81224, AAC81225, AAC81262, AAC81263). The HD-PTP gene is located on chromosome 3p21.3. This region is frequently found to be deleted in lung cancers, and is therefore thought to contain a tumour suppressor gene. The invention also relates to expression vectors and host cells containing human HD-PTP nucleic acids; the recombinant production of HD-PTP; anticancer drugs containing HD-PTP; gene therapy compositions containing DNA encoding HD-PTP; diagnostic reagents containing HD-PTP oligonucleotides; antibodies specific for HD-PTP; and an immunoassay method using HD-PTP-specific antibodies for use in cancer diagnosis and investigation. HD-PTP proteins, nucleic acids and antibodies may be used in the treatment, investigation and diagnosis of cancers, particularly those of the lung. The present sequence represents the full-length human HD-PTP.
XX SQ Sequence 1636 AA;

Query Match          3.7%; Score 118.5; DB 21; Length 1636;
Best Local Similarity   21.8%; Pred. No. 0.24;
Matches 138; Conservative 73; Mismatches 224; Indels 197; Gaps 34

QY    24 AAEIGREEDPLVLLSL-----VLG-----FVEHFLAVNRVIPTWPEL-TFQP- 65
       | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db    283 alkklggqdtvgdalfmtdyigdkynsakkdndfiyh-----eavpalddtqp 333
                                         |||||

QY    66 --SPADDP--PGGLTFYPVADLSIIIAALYAARETAQIRGAVDLVIPREGGVSSRELVKKVS 122
       | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db    334 kgaplvkplvnpdpdvtpgdifaklvpmaahes-----sllyseekallremmakie 398
                                         |||||

QY    123 DVIWNSLSRSYFKDRAHQISLFSTGTGKLDSGGV----AFA-----VGCAQAL----- 168
       | |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db    389 d-----knveildafmdsmaldbetvdnldayshlpqdimekcgaalsvrpd 433

```

Qy	169	-----GLRDVHLALSEDHAWVVFPGNGEQTAETVYTWICKGNEDRR	
Db	434	tvrnlgsmqvlsgvftdv easikdirldleed-----elleqktfeavgdgag	483
Qy	208	GQTVNAGVAE--RSWLKGSYRCDRMEVAFMVCAINPSIDLHTDLSLELLQLOKLLW	265
Db	484	sitskaelaevrrew----akymevhek---astfnseihramlnlhvgnrlrls-----	530
Qy	266	LLYDLGHLERYPMALGNLADLELEPTP--GRDPDLTYH-KGIASAKTVYRDEHIVPYM	322
Db	531	-----gpldqvraal-----ptpalspedkavlnqlkilakvqemrdqrv----	571
Qy	323	YLACYHCRNRNVREALQAWDATVI-QDYNVCRE-DEEYKKEFFEVDNDVIPNLKE--	378
Db	572	-----sleqqlldliqdditlaslvttdhsemkklfeeqllkkydqllkvytleqnlaaqdr	625
Qy	379	AASLLAGEERPEQ---SQGTQSQSALQ----DPECFAHLLRFVDGICKWEESPTP	430
Db	626	vlcalteanqvaavrvlsldqgkwnstlqtlvasyeyadedlmk-----ksgegrdfy	679
Qy	431	V-LHVGWATFLVQSLGRFEGQVRQKVRIVSR-----EAEAAEAE	468
Db	680	adleskvaallertqstcqareaaqqldrelkktpprptapklprrseeseavsg	739
Qy	469	EPWGEERE---GRRRGPR-----RESKPEEPPPPKKPALDKLGTGQGAVSGPPRKP	519
Db	740	dp-peelrsipppdmvagpripdtflgsatplhfppspfps---stgggphyisgp--lpp	793
Qy	520	GTVAGTARGEGGSTAQVPAPAASPPPEGPVL	551
Db	794	gtyS-----gptqlqlpragphampvapgpal	821
RESULT 8			
Id	AAW92296	AAW92296 standard; peptide; 1301 AA.	
AC	AAW92296;		
DT	28-APR-1999	(first entry)	
DE	Human alpha-1 (XVIII) collagen chain common sequence HU18(common)36.		
KW	Human; type XVIII collagen; liver disease; cirrhosis; detection;		
KW	hepatocellular carcinoma; diagnosis.		
OS	Homo sapiens.		
PN	W09856399-A1.		
XX	17-DEC-1998.		
XX	12-JUN-1998; 98WO-US12327.		
PR	12-JUN-1997; 97US-0049369.		
XX	(FIFI-) ACAD FINLAND.		
PA	(FIBR-) FIBROGEN INC.		
PA	(INRM) INST NAT SANTE & RECH MEDICALE.		
PI	Clement B, Pihlajaniemi T, Rehn M;		
XX	WPI; 1999-070292/06.		
PT	Diagnosis and monitoring of liver disease by measuring collagen type		
PT	XVIII levels - with elevated levels indicative of disease,		
PT	especially cirrhosis or hepatocellular carcinoma		
XX	Example 6; Fig 8; 56pp; English.		
PS	A method has been developed for the detecting liver disease. The metho		
CC	comprises: (a) reacting a patient sample with antibodies(Ab) specifi		

AC AAW50192;
 XX 28-JUL-1998 (first entry)
 XX Amino acid sequence of salivary protein CON-1.
 DE Salivary glycoprotein; CON-1; CON-2; alpha-glucosidase inhibitor;
 XX retrovirus; inhibition; diabetes; HIV.
 KW Homo sapiens.
 XX WO9809981-A1.
 XX 12-MAR-1998.
 XX 08-SEP-1997; 97WO-US15799.
 PF 09-SEP-1996; 96US-0024712.
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX Azen EA, Pan D;
 XX WPI; 1998-193547/17.
 DR N-PSDB; AAV22062.
 XX Isolated salivary glyco-protein CON-1 and CON-2 compositions - which
 XX have alpha-glucosidase inhibitory activity, useful for treating
 PT diabetes or retrovirus, particularly HIV infection
 XX Disclosure; Page 32; 54pp; English.
 XX This amino acid sequence is of the salivary glycoprotein CON-1,
 CC apotent alpha-glucosidase inhibitor, which is useful in preventing
 CC cellular penetration of retroviruses. In the method of the invention
 CC CON-1 inhibits alpha-glucosidase processing of the retroviruses
 CC required for proper engagement of the virion with its cellular
 CC receptor. The salivary glycoproteins CON-1 and CON-2 and derivatives,
 CC have AGS inhibitory activity and can be used to treat patients with
 CC diabetes or patients infected with retroviruses such as HIV.
 XX Sequence 124 AA;
 SQ

Query Match 3.7%; Score 117.5; DB 19; Length 124;
 Best Local Similarity 32.3%; Pred. No. 0.0053;
 Matches 31; Conservative 9; Mismatches, 25; Indels 31; Gaps 5;
 Qy 469 EPWGEAREGRRRRPRRE--SKPEPPPPKPKALDKGLGTGGQAVSGPP-----R 516
 Db 16 qpqgppppgkpgppgpgggnkpgppppgk-----qgpppggdnksqsr 62
 Qy 517 KPGCTVAGTARGEGGTAQVPAP----AASPPPEG 548
 Db 63 sppgkpgpp--pdggnqgpgpppppgkpgpppgq 96

RESULT 11
 AAY06309
 ID AAY06309 standard; Protein; 1566 AA.
 XX AC AAY06309;
 XX 06-SEP-1999 (first entry)
 XX Human p53 regulatory protein RB18A.
 XX RB18A; p53 regulatory protein; apoptosis; neoplasia; inflammation;
 KW wound healing; graft rejection; reperfusion injury;
 KW myocardial infarction; stroke; traumatic brain injury;
 KW neurodegenerative disease; ischaemia; toxemia; infection; AIDS;
 KW hepatitis; breast cancer; ovarian cancer; colon cancer; diagnosis;
 KW therapy; human.

XX OS Homo sapiens.
 XX Key
 FH Domain Location/Qualifiers
 FT /note= "DNA binding domain; this polypeptide is
 FT specifically claimed in Claim 4"
 FT 1234..1406
 FT /note= "p53 binding and homo-oligomerisation
 FT domains; this polypeptide is specifically
 FT claimed in Claim 4"
 FT 955..958
 FT /note= "nuclear localisation signal"
 FT 979..982
 FT /note= "nuclear localisation signal"
 FT 996..999
 FT /note= "nuclear localisation signal"
 FT 1483..1493
 FT /note= "nuclear localisation signal" ;
 FT 436..1566
 FT /note= "this polypeptide is specifically claimed
 FT in Claim 4"
 FT 436..1228
 FT /note= "this polypeptide is specifically claimed
 FT in Claim 4"
 FT 436..927
 FT /note= "this polypeptide is specifically claimed
 FT in Claim 4"
 FT 1537..1566
 FT /note= "this polypeptide is specifically claimed
 FT in Claim 4"
 FT 1234..1566
 FT /note= "this polypeptide is specifically claimed
 FT in Claim 4"
 XX WO9931231-A1.
 XX 24-JUN-1999.
 XX 14-DEC-1998; 98WO-EP08560.
 XX 15-DEC-1997; 97EP-0403051.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Frade R;
 XX WPI; 1999-395177/33.
 XX N-PSDB; AAX59124.
 XX New p53 regulatory protein (RB18A) useful as, e.g. sources of probes
 XX and primers to detect the transcription rate and abundance of RB18A
 XX mRNA in lymphocytes
 XX Claim 5; Fig 3; 87pp; English.
 XX This sequence represents human RB18A (recognized by PAb1801
 CC antibody), a novel 205 kDa p53 regulatory protein that inhibits
 CC p53-induced apoptosis. RB18A cDNA (see AAX59124) was isolated from
 CC a human heart cDNA library. RB18A shares some antigenic epitopes
 CC with p53. It regulates the sequence-specific DNA binding function
 CC of p53. Its C-terminal domain induces a dose-dependent
 CC reassociation of DNA double strands, stimulates in vitro the
 CC sequence-specific DNA binding of p53, and functionally interacts
 CC and stimulates in vivo the transactivation activity of p53. RB18A
 CC polypeptides can be used to block the G1 phase of the cell cycle
 CC and/or regulate apoptosis, and so control cell growth. This makes
 CC them good candidates for antineoplastic therapy. Pharmaceutical
 CC compositions comprising RB18A polypeptides or polynucleotides are
 CC useful for preventing or treating a variety of human and veterinary
 CC diseases, e.g. neoplasia, inflammation, wound healing, graft
 CC rejection, reperfusion injury, myocardial infarction, stroke,
 CC traumatic brain injury, neurodegenerative diseases, ageing,

CC ischaemia, toxemia, infection, AIDS and hepatitis. Preferably,
 CC they are used as antineoplastic compositions, or compositions
 CC directed against any other cell proliferation disease, especially
 CC in the treatment of colon, breast or ovarian cancer.
 XX
 SQ Sequence 1566 AA;

Query Match 3.7%; Score 117.5; DB 20; Length 1566;
 Best Local Similarity 22.4%; Pred. No. 0.27;
 Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDNYNCREDEEYKEFEFVANDVIPNLKEAASLLEAGEERPEGEOSQ----- 395
 Db 820 ytdpadliadaagssdsptnhfhdvdfnpdll-nsgsqsgfgeeyfdesqsgdnd 878

QY 396 ---GTOSQ-----GSALQDPECFALLRF 416
 Db 879 dfkgfasqalntlgvmlgdngetkfgnngadtvdfsiisvagkalapadimeh---- 934

QY 417 YDGICKWEGSPPTVLHVGMATFLVQSLGRFEGQVRQKRVISREAEAEPEWGEAR 476
 Db 935 -----hsgsgpllttg-----dlgkektqkrvk---egnsgtsnstlsgp-glisk 976

QY 477 EGRR-RGPRRESKPEPPPPKPPALDKGLGTGGAVSGPPRPKPPGTAGTARPEGSTA 535
 Db 977 pgkrstpsndgskdkppkrkkadtekgspshs-snrpftpp-tstgsgkspgsagrs 1034

QY 536 QVPAPASPP-----PEGPVLTFQSEKMGKMKELLVATKINSSAIK-----LQL 579
 Db 1035 qtpgvgatppipkitiqikgtvmv---gkpsshsqtstsgsvssgskshshsssss 1091

QY 580 TAQSQVMKKQKVSTPSTDTLS 601
 Db 1092 sastsgkmkssksegssskls 1113

RESULT 12
 AAY69669
 ID AAY69669 standard; Protein; 1581 AA.
 AC AAY69669;
 XX
 DT 08-MAY-2000 (first entry)
 XX Human thyroid receptor-associated protein TRAP220.
 XX
 KW Thyroid receptor-associated protein; TRAP220; TRAP complex; coactivator;
 KW nuclear hormone receptor; thyroid receptor; vitamin D receptor;
 KW oestrogen receptor; mineralcorticoid receptor;
 KW peroxisome proliferation-activated receptor; LXXLL motif; drug screening;
 KW detection.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 604..608
 FT /note= "LXXLL motif 1"
 FT Region 645..649
 FT /note= "LXXLL motif 2"
 XX
 PN WO200001820-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 01-JUL-1999; 99WO-US15052.
 XX
 PR 06-JUL-1998; 98US-0110517.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 XX
 PI Roeder RG, Fondell JD, Xingyuan C, Ito M;
 XX

DR WPI: 2000-147418/13.
 DR N-PSDB; AA287101.
 XX
 PT New isolated Thyroid Receptor-Associated Proteins which act as nuclear
 PT hormone receptor coactivators used for identifying modulators of
 PT hormones or nuclear hormone receptors
 XX
 PS Claim 9; Fig 5A; 114pp; English.
 XX
 CC The present sequence represents human thyroid receptor-associated
 CC protein TRAP220. TRAP220 is a member of a complex of TRAPs which
 CC act as coactivators for nuclear hormone receptors, binding
 CC to such receptors in a ligand-dependent manner and are required for
 CC functional interactions between the receptor and genes whose
 CC transcription is regulated by these receptors. Nuclear hormone receptors
 CC include thyroid receptors (TRs), vitamin D receptors (VDRs), oestrogen
 CC receptors (ERs), mineralcorticoid receptors (MRs) and peroxisome
 CC proliferation-activated receptors (PPARs). TRAP220 contains two of the
 CC LXXLL motifs that have been implicated in nuclear hormone receptor-
 CC coactivator interactions. TRAP220, and a second protein of the
 CC invention, TRAP100 (AAY69670), and nucleotides encoding these proteins
 CC may be used to modulate the activity of a nuclear hormone receptor, or to
 CC screen for agents that modulate receptor or hormone activity. Proteins,
 CC nucleic acids and antibodies may also be used therapeutically and for
 CC detection of TRAP220 and TRAP100 or their associated nucleotides.
 XX
 SQ Sequence 1581 AA;

Query Match 3.7%; Score 117.5; DB 21; Length 1581;
 Best Local Similarity 22.4%; Pred. No. 0.28;
 Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDNYNCREDEEYKEFEFVANDVIPNLKEAASLLEAGEERPEGEOSQ----- 395
 Db 835 ytdpadliadaagssdsptnhfhdvdfnpdll-nsgsqsgfgeeyfdesqsgdnd 893

QY 396 ---GTOSQ-----GSALQDPECFALLRF 416
 Db 894 dfkgfasqalntlgvmlgdngetkfgnngadtvdfsiisvagkalapadimeh---- 949

QY 417 YDGICKWEGSPPTVLHVGMATFLVQSLGRFEGQVRQKRVISREAEAEPEWGEAR 476
 Db 950 -----hsgsgpllttg-----dlgkektqkrvk---egnsgtsnstlsgp-glisk 991

QY 477 EGRR-RGPRRESKPEPPPPKPPALDKGLGTGGAVSGPPRPKPPGTAGTARPEGSTA 535
 Db 992 pgkrstpsndgskdkppkrkkadtekgspshs-snrpftpp-tstgsgkspgsagrs 1049

QY 536 QVPAPASPP-----PEGPVLTFQSEKMGKMKELLVATKINSSAIK-----LQL 579
 Db 1050 qtpgvgatppipkitiqikgtvmv---gkpsshsqtstsgsvssgskshshsssss 1106

QY 580 TAQSQVMKKQKVSTPSTDTLS 601
 Db 1107 sastsgkmkssksegssskls 1128

RESULT 13
 AAY69671
 ID AAY69671 standard; Protein; 1581 AA.
 XX
 AC AAY69671;
 XX
 DT 08-MAY-2000 (first entry)
 XX Human thyroid receptor-associated protein TRAP220 mutant a.
 XX
 KW Thyroid receptor-associated protein; TRAP220; TRAP100; coactivator;
 KW TRAP complex; nuclear hormone receptor; thyroid receptor;
 KW vitamin D receptor; oestrogen receptor; mineralcorticoid receptor;
 KW peroxisome proliferation-activated receptor; LXXLL motif; drug screening;
 KW detection; mutant; mutein.
 KW

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 03:45:17 ; Search time 8321.72 Seconds
(without alignments)
10427.796 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGCTCTGTGAACCTCTGGCC.....AGCCACGGGCGCGCCGCCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_est4:*
 - 5: gb_est5:*
 - 6: gb_est6:*
 - 7: gb_est7:*
 - 8: gb_est8:*
 - 9: gb_est9:*
 - 10: gb_est10:*
 - 11: gb_est11:*
 - 12: gb_est12:*
 - 13: gb_est13:*
 - 14: gb_est14:*
 - 15: gb_est15:*
 - 16: gb_est16:*
 - 17: gb_est17:*
 - 18: gb_est18:*
 - 19: gb_est19:*
 - 20: gb_est20:*
 - 21: gb_est21:*
 - 22: gb_est22:*
 - 23: gb_est23:*
 - 24: gb_est24:*
 - 25: gb_est33:*
 - 26: gb_est34:*
 - 27: gb_est35:*
 - 28: gb_est36:*
 - 29: gb_est37:*
 - 30: gb_est38:*
 - 31: gb_est39:*
 - 32: gb_est40:*
 - 33: em_estba:*
 - 34: em_estfun:*
 - 35: em_esthum1:*
 - 36: em_esthum2:*
 - 37: em_esthum3:*
 - 38: em_esthum4:*
 - 39: em_esthum5:*
 - 40: em_esthum6:*
 - 41: em_esthum7:*
 - 42: em_esthum8:*
 - 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estini:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estp11:*
- 73: em_estp12:*
- 74: em_estp13:*
- 75: em_estp14:*
- 76: em_estp15:*
- 77: em_estp16:*
- 78: em_estp17:*
- 79: em_estp18:*
- 80: em_estp19:*
- 81: em_estp110:*
- 82: em_estrol:*
- 83: em_estrol2:*
- 84: em_estrol3:*
- 85: em_estrol4:*
- 86: em_estrol5:*
- 87: em_estrol6:*
- 88: em_estrol7:*
- 89: em_estrol8:*
- 90: em_estrol9:*
- 91: em_estrol10:*
- 92: em_estrol11:*
- 93: em_estrol12:*
- 94: em_estrol13:*
- 95: em_estrol14:*
- 96: em_estrol15:*
- 97: em_estrol16:*
- 98: em_estrol17:*
- 99: em_estrol18:*
- 100: em_estrol19:*
- 101: em_estrol20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	809.2	8.8	912	106	AL582775	AL582775 AL582775
C 2	773.6	8.4	865	106	AL560096	AL560096 AL560096
C 3	688	7.5	703	105	AL045547	AL045547 DKEP434L
C 4	658.2	7.2	898	105	AL521542	AL521542 AL521542
C 5	613.4	6.7	680	165	BE267789	BE267789 601124773
C 6	574.4	6.3	792	106	AL561613	AL561613 AL561613
C 7	566.8	6.2	641	122	AW950638	AW950638 EST362708
C 8	560.4	6.1	580	165	BE267618	BE267618 601124837
C 9	532	5.8	574	122	AW968425	AW968425 EST380501
C 10	489.6	5.3	799	150	BF526328	BF526328 602071372
C 11	484	5.3	523	169	BF803247	BF803247 IL5-C1014
C 12	474.4	5.2	876	105	AL521543	AL521543 AL521543
C 13	471.4	5.1	897	106	AL560097	AL560097 AL560097
C 14	466.8	5.1	525	11	AA776738	AA776738 ah9f08.s
C 15	461.8	5.0	525	19	AI333944	AI333944 q27b09.x
C 16	460.8	5.0	533	11	AA743431	AA743431 ny23f01.s
C 17	460.4	5.0	630	165	BE281439	BE281439 601154960
C 18	459	5.0	503	15	AI082242	AI082242 ox79g05.x
C 19	449.4	4.9	497	104	AI953167	AI953167 wp99d02.x
C 20	447.8	4.9	496	116	AW439193	AW439193 xt16c10.x
C 21	445	4.8	490	152	BG328937	BG328937 602428083
C 22	444	4.8	511	10	AA705195	AA705195 zj96e04.s
C 23	440.2	4.8	500	148	BF448417	BF448417 7n8e07.x
C 24	440	4.8	496	145	BF195671	BF195671 7n8e07.x
C 25	440	4.8	507	112	AW150068	AW150068 x948b02.x
C 26	434.2	4.7	738	175	BG251389	BG251389 602364035
C 27	431	4.7	486	145	BF195544	BF195544 7n89f07.x
C 28	427	4.7	687	152	BG326214	BG326214 602424925
C 29	425.4	4.6	591	112	AW161381	AW161381 aub1a12.y
C 30	422	4.6	467	14	AA973790	AA973790 oa45904.s
C 31	420	4.6	464	22	AI633641	AI633641 th70f01.x
C 32	418.8	4.6	467	191	Z98467	Z98467 HSZ98467 DK
C 33	418.4	4.6	464	18	AI288914	AI288914 q191g06.x
C 34	417.4	4.5	474	17	AI224561	AI224561 q134d04.x
C 35	411.4	4.5	856	153	BG422364	BG422364 602446924
C 36	405	4.4	449	18	AI275473	AI275473 q174a11.x
C 37	404	4.4	437	223	AQ053443	AQ053443 RPC111-51
C 38	402.4	4.4	448	19	AI332627	AI332627 qq28b05.x
C 39	402.4	4.4	460	19	AI371104	AI371104 ta30a02.x
C 40	402	4.4	446	19	AI371103	AI371103 ta30a01.x
C 41	402	4.4	480	9	AA604513	AA604513 no73c10.s
C 42	399.8	4.4	448	10	AA636030	AA636030 nr37h07.s
C 43	394.8	4.3	808	154	BG476481	BG476481 602522015
C 44	394.6	4.3	443	24	AI738671	AI738671 w11g09.x
C 45	392	4.3	437	15	AI056630	AI056630 oz18e01.x

ALIGNMENTS

RESULT 1	AL582775/c	AL582775	912 bp	mRNA	EST	16-FEB-2001
LOCUS	AL582775	LTI_NFL010_BC2	Homo sapiens	cdna clone	CS0DL011Y120	3
DEFINITION	prime, mRNA sequence.					
ACCESSION	AL582775					
VERSION	AL582775.1	G1:12951093				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.					
TITLE	Full-length cDNA libraries and normalization					
JOURNAL	Unpublished (2001)					
COMMENT	Contact: Genoscope					

Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		Location/Qualifiers	
FEATURES		source	
		1..912	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="CS0DL011Y120"	
		/clone_lib="LTI_NFL010_BC2"	
		/sex="male"	
		/tissue_type="B cells from Burkitt lymphoma"	
		/notes="Vector: pCMVSPORT 6; Site: 1. NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"	
BASE COUNT		179 a	178 c 307 g 214 t 34 others
ORIGIN			
Query Match		8.8%; Score 809.2; DB 106; Length 912;	
Best Local Similarity		94.5%; Pred. No. 2.3e-64;	
Matches		867; Conservative 29; Mismatches 15; Indels 6; Gaps 6;	
QY	7925	GGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCACTCAGGCACAGTCGCAAGTGCA	7984
Db	912	GGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCACTCAGGCACAGTCGCAAGTGCA	853
QY	7985	GATGAAGAAGCAGAAAGTGCCACCCCTAGTACTGACTGCTGCTTCTTCCCAAGCGCA	8044
Db	852	GATGAAGAAGCAG-AAARKTTCCACCCCTAGTACTGACTGCTGCTTCTTCCCAAGCGCA	794
QY	8045	GGCCAAAGGGCTCTGAACTACTTGGGGACTTTCGGACCGCTTGTGGGGACCCAGGCTCCG-C	8103
Db	793	SCSAAAGSCTCTGAACTACTTGGGGACTTTCGGACCGCTTGTGGGGACCCAGGCTCCG-C	734
QY	8104	CTTAGTCCCCCACTCTGAGCGCCATGTTCTGCCGCCCAAGGGGACAGGCTCCACC	8163
Db	733	CTTAGTCCCCCACTCTGAGCGCCATGTTCTGCCGCCCAAGGGGACAGGCTCCACC	674
QY	8164	TCATACCAAAACCTTAGTCTCCCGTCCCGAGTACAGTCTGTATCAAAACCCACGATTCT	8223
Db	673	TCATACCAAAACCTTAGTCTCCCGTCCCGAGTACAGTCTGTATCAAAACCCACGATTCT	614
QY	8224	CCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGTAATATAGGTCTCTTCTCCCAAGAT	8283
Db	613	CCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGTAATATAGGTCTCTTCTCCCAAGAT	554
QY	8284	CCAGCGGGCCAAATGGAAACCTCAGCTGGGTCTTAATACAGTCTTTAAAGGCCACG	8343
Db	553	CCAGCGGGCCAAATGGAAACCTCAGCTGGGTCTTAATACAGTCTTTAAAGGCCACG	494
QY	8344	CCCTAGAAACCCCAAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAACGTTACTCA	8403
Db	493	CCCTAGAAACCCCAAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAACGTTACTCA	434
QY	8404	GGGTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCCAAGGAGCAGAGTCC	8463
Db	433	GGGTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCCAAGGAGCAGAGTCC	375
QY	8464	CACCTCAGCGCCGGGAGCCCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACTTCA	8523
Db	374	CACCTCAGCGCCGGGAGCCCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACTTCA	316
QY	8524	GAATTCATCTTTCGGAATCCCAAGCTCCCTGCCCAAAATAACTTCAGTCTGCTCCAG	8583
Db	315	GAATTCATCTTTCGGAATCCCAAGCTCCCTGCCCAAAATAACTTCAGTCTGCTCCAG	257

```

QY 8584 AATTGGAAATCCTAGTTTCTCTCTCTCGTATCCGAGTCTGGGACACAAACATCCGCGC 8643
Db 256 AATTGGAAATCCTAGTTTCTCTCTCTCGTATCCGAGTCTGGGACACAAACATCCGCGC 197
QY 8644 CCAGCGTATGAGCATCTGAGCGCGCGCTCTTCTGACGAACTGCGCCCGGATCAGA 8703
Db 196 CCACCGCTATGAGCATCTGAGCGCGCGCTCTTCTGACGAACTGCGCCCGGATCAGA 137
QY 8704 GCAGGACCTCCCTCCGACCCCTCTGGAACTCTCCAGAGGTCCAGCCCATCTCGGAGCAT 8763
Db 136 GCAGGACCTCCCTCCGACCCCTCTGGAACTCTCCAGAGGTCCAGCCCATCTCGGAGCAT 77
QY 8764 CCCGAGGAAATCTCGAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTCTGTT 8823
Db 76 CCCGAGGAAATCTCGAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTCTGTT 18
QY 8824 TTGTACATAGATTATT 8840
Db 17 TAGACATAGATTATT 1

RESULT 2
AL560096/C 865 bp mRNA EST 16-FEB-2001
LOCUS AL560096 LTI_FL011_Bc1 Homo sapiens cDNA clone CS0DG002YG12 3 prime
DEFINITION , mRNA sequence.
ACCESSION AL560096
VERSION AL560096.1 GI:12906228
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li-W.B., Gruber C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YG12"
/clone.lib="LTI_FL011_Bc1"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/notes="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 165 a 183 c 306 g 204 t 7 others
ORIGIN

Query Match 8.4%; Score 773.6; DB 106; Length 865;
Best Local Similarity 97.1%; Pred. No. 3.7e-61;
Matches 845; Conservative 5; Mismatches 14; Indels 6; Gaps 6;

QY 7817 AGGTGGCAGCAGCGTCTAGGTCCAGCACCCG-CAGCATCACACCGCGGAGGTCCAG 7875
Db 865 AAGTGCAGCAGCGTCTAGGTCCAGCACCCGCGGAGGTCCAG 806
QY 7876 TCCTCACTTCCAGAGTGAAGATCAAGGCGATCAAGGAGTCTGTGGCCACCAAGA 7935
Db 1

```

```

Db 805 TKCTCACTTTCAGAGTGAGAAAGATGAAGGCGATGAAGGAGCTGCT- GTGGCCACCAAGA 747
QY 7936 TCAACTCGAGCGCCCATCAAGCTGCAACTCACGGCACAGTCGAAGTGCAGATCAAGAAGC 7995
Db 746 TCAACTCGAGCGCCCATCAAGCTGCAACTCACGGCACAGTCGAAGTGCAGATCAAGAAGC 687
QY 7996 AGAAAGTGTCCACCCCTAGTGTACTACTCTCTCTTCTCAAGCGGACAGCCAAAGGCC 8055
Db 686 AGAAAGTGTCCACCCCTAGTGTACTACTCTCTTCTTCTCAAGCGGACAGCCAAAGGCC 627
QY 8056 TCTGAACTACTGGGACACTTCGGACCGCTGTGGGACCCAGGCTCGCGCTTCTAGTCCCA 8115
Db 626 TCTGAACTACTGGGACACTTCGGACCGCTGTGGGACCCAGTTCGCCCTTCTAGTCCCA 567
QY 8116 ACTCTGAGCGCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGCTCACCTCTACCCAAAC 8175
Db 566 ACTCTGAGCGCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGCTCACCTCTACCCAAAC 507
QY 8176 CTAGTTTCCCGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCCAGCTCAGAA 8235
Db 506 CTAGTTTCCCGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCCAGCTCAGAA 447
QY 8236 CCAGGCGCTCGCCCGAGTGTAGATATAGTCTCTTCTCCAGAACTCCAGCGGCCA 8295
Db 446 CCA-GGCTCTGCCCGAGTGTAGATATAGTCTCTTCTCCAGAACTCCAGCGGCCA 388
QY 8296 ATGGAACCTCACGCTGGGTCTTAATTACAGTCTTTAAAGGCCAGCCCTAGAAACCC 8355
Db 387 ATGGAACCTCACGCTGGGTCTTAATTACAGTCTTTAAAGGCCAGCCCTAGAAACCC 328
QY 8356 AAGCTCTCTCGGAAACCGCTCACCTAGAGCAGACCAACGTTTACTTCAGGGTCTCTCCA 8415
Db 327 AAGCTCTCTCGGAAACCGCTCACCTAGAGCAGACCAACGTTTACTTCAGGGTCTCTCCA 268
QY 8416 GCTTCTAGAGCTGAGTTTACCCCTTACCCCAAGGGGACAGGTCACACCTCCAGCCC 8475
Db 267 GCTTCTAGAGCTGAGTTTACCCCTTACCCCAAGGGGACAGGTCACACCTCCAGCCC 209
QY 8476 GGGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCCGCACCTTCAGAAATTCATATC 8535
Db 208 GGGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCCGCACCTTCAGAAATTCATATC 150
QY 8536 TTGCGAATCCAAAGTCCCTGCGCCCAAAATAACTTCTAGTCTCTGTTCCAGAAATTTGGAATC 8595
Db 149 TTGCGAATCCAAAGTCCCTGCGCCCAAAATAACTTCTAGTCTCTGTTCCAGAAATTTGGAATC 91
QY 8596 CTAGTTTCTCTCTCTGTTATCCCGAGTCTGGGACACAAACTCCGCCCGGAGCTTATGA 8655
Db 90 CTAGTTTCTCTCTCTGTTATCCCGAGTCTGGGACACAAACTCCGCCCGGAGCTTATGA 31
QY 8656 GCATCTGAGCGCCCGCTCTTCTGACGA 8685
Db 30 GCATCTGAGCGCCCGCTCTTCTGACGA 1

RESULT 3
AL045547 703 bp mRNA EST 29-FEB-2000
LOCUS AL045547
DEFINITION DK22P34L125_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045547
VERSION AL045547.1 GI:5433678
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 703)
JOURNAL Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann
COMMENT Unpublished (1999)
Contact: Duesterhoeft A

```


MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKF2); Email s.wiemann@dkfz-heidelberg.de;
 consortium of Otagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 NO 5' sequence available.

This clone (DKF2p434L125) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
 1..703
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKF2p434L125"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
 136 a 227 c 185 g 151 t 4 others

BASE COUNT 136 a 227 c 185 g 151 t 4 others
 ORIGIN
 Query Match 7.5%; Score 688; DB 105; Length 703;
 Best Local Similarity 99.4%; Pred. No. 1.9e-53;
 Matches 699; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 5720 GCCATTGGCTGGCTGGAAGTCTTTGGTATGTGTAGAGAGTGTCTGAGAAGAGAA 5779
 Db 1 GCCATTGGCTGGCTGGAAGTCTTTGGTATGTGTAGAGAGTGTCTGAGAAGAGAA 60
 Qy 5780 GGGCCCTGAGCTCGAGGGGAGGCCACCCCTCGAGTCTGCCCCAGGCTCAGCCAGCA 5839
 Db 61 GGGCCCTGAGCTCGAGGGGAGGCCACCCCTCGAGTCTGCCCCAGGCTCAGCCAGCA 120
 Qy 5840 GTCTGTAGAGCCAGGAGGAGACAGGTAGAAGGGCTGGCAGGAGTGGAGGTGGGAGT 5899
 Db 121 GTCTGTAGAGCCAGGAGGAGACAGGTAGAAGGGCTGGCAGGAGTGGAGGTGGGAGT 180
 Qy 5900 GGAGATGGAGAGACTCCCTGGGATCTTCCTGTGGCCCTTCTGGGTGTCCTGGTGGG 5959
 Db 181 GGAGATGGAGAGACTCCCTGGGATCTTCCTGTGGCCCTTCTGGGTGTCCTGGTGGG 240
 Qy 5960 GCATTGTGCGCAGGAGGAGCTGGGGCTGCTCCCTGAGGATCTCTGCTCACTCCCA 6019
 Db 241 GCATTGTGCGCAGGAGGAGCTGGGGCTGCTCCCTGAGGATCTCTGCTCACTCCCA 300
 Qy 6020 TCCAGGGCATTGCTCAGCCAGACCTACTATCGGGATGAACACATCTACCCCTACATGT 6079
 Db 301 TCCAGGGCATTGCTCAGCCAGACCTACTATCGGGATGAACACATCTACCCCTACATGT 360
 Qy 6080 ACCTGGCTGGTACACTGTGCAACCGCAATGTGGGGAAGCCCTCGAGGCTGGGGGG 6139
 Db 361 ACCTGGCTGGTACACTGTGCAACCGCAATGTGGGGAAGCCCTCGAGGCTGGGGGG 420
 Qy 6140 ACAGGGCACTGTCTATCCAGAGTGAAGTCCCCCTACTAGGGCTCGCAGGCTGTCTTT 6199
 Db 421 ACAGGGCACTGTCTATCCAGAGTGAAGTCCCCCTACTAGGGCTCGCAGGCTGTCTTT 480
 Qy 6200 CTTCCTCCCTCCATCAGTTTCCAAACCACTCGTCCAGGACTGAGGCTGGCTTCCACGCC 6259
 Db 481 CTTCCTCCCTCCATCAGTTTCCAAACCACTCGTCCAGGACTGAGGCTGGCTTCCACGCC 540
 Qy 6260 CATCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCATTTACCA-GGAGGTAGGAC 6318
 Db 541 CATCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCATTTACCAAGGAGTANGAC 600
 Qy 6319 CCTGATTAAGGTGCACATCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 6378
 Db 601 CCTGATTAAGGTGCACATCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 660

Qy 6379 GAACAGTCTCAATCTCAATGTTTAAACCACCATCATCCAGCA 6421
 Db 661 GAACAGTCTCAATCTCAATGTTTAAACCACCATCATCCANCA 703

RESULT 4

AL521542/c

LOCUS

DEFINITION

AL521542 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB003YA04 3

prime, mRNA sequence.

ACCESSION

AL521542

VERSION

AL521542.1 GI:12785035

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 898)

AUTHORS

Li, W.B., Gruber, C., Jesses, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..898

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DB003YA04"

/clone_lib="LTI_NFL004_NBC2"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact: Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax: (1) 301 610

8371 Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 190 a 188 c 297 g 206 t 17 others

ORIGIN

Query Match 7.2%; Score 658.2; DB 105; Length 898;

Best Local Similarity 89.0%; Pred. No. 7.6e-51;

Matches 803; Conservative 8; Mismatches 68; Indels 23; Gaps 9;

Qy 7978 AAGTGCAGATGAAGAGAGAGAGTGTCCACCCCTAGTACTACTACTGTCTTTCTCA 8037

Db 898 AAGTGSAGATGAAGAGAGAGAGTGTCCACCCCTAGTACTACTACTGTCTTTCTCA 839

Qy 8038 AGCGGCGAGCAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAG 8097

Db 838 AGCGGCGAGCAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAG 779

Qy 8098 CTCGCCCTTAGTCCCCCAACTCTGAGCCCATGTTCCTGCCCCAGCCCAAGGGAGGAC 8157

Db 778 CTCGCCCTTAGTCCCCCAACTCTGAGCCCATGTTCCTGCCCCAGCCCAAGGGAGGAC 719

Qy 8158 CTCACCTTACCCAAACCCCTAGGTTCGGGTCCCGAGTACAGTCTGTATCAACCCAGCA 8217

Db 718 CTCACCTTACCCAAACCCCTAGGTTCGGGTCCCGAGTACAGTCTGTATCAACCCAGCA 659

Qy 8218 TTTTCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCCGTTAGATATAGTCTCTTTCTCC 8277

Db 658 TTTTCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCCGTTAGATATAGTCTCTTTCTCC 600

Qy 8278 CAGATCCAGCCGCCCAATGGAACCTCAGGCTGGGTCTTAATTTACCAGTCTTTAAAGG 8337

DEFINITION AL561613 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL011Y120 5
prime, mRNA sequence.
ACCESSION AL561613
VERSION AL561613.1 GI:12909216
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL011Y120"
/clone_lib="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 119 a 259 c 252 g 157 t 5 others
ORIGIN
Query Match 6.3%; Score 574.4; DB 106; Length 792;
Best Local Similarity 84.0%; Pred. No. 2.6e-43;
Matches 739; Conservative 4; Mismatches 0; Indels 137; Gaps 3;
Qy 1854 GGTCCTAGTGTGGATGTAAGCGCGAGGTGGCGAGGGGACCGAGGCGAGACTCTC 1913
Db 1 GGTCCCTAGTGTGGATGTAAGCGCGAGGTGGCGAGGGGACCGAGGCGAGACTCTC 60
Qy 1914 CTTGGGGTTTGGGGCTTGACCTGGTGGCTTTCTGGACAGACTTTACAGCCCCGGGG 1973
Db 61 CTTGGGGTTTGGGGCTTGACCTGGTGGCTTTCTGGACAGACTTTACAG-CCCCGGGG 119
Qy 1974 GCACAGTCGTAGAGAGGGGGGGCGCCATTGGGGCTCTCATTTGGGTGCTTGGGGC 2033
Db 120 GCACAGTCGTAGAGAGGGGGGGGGCGCCATTGGGGCTCTCATTTGGGTGCTTGGGGC 179
Qy 2034 GCACCCCATCGGTACCGGGGCTCCCGAATTGTGGGGGACAAAAGCTCTGCAGTCTC 2093
Db 180 GCACCCCATCGGTACCGGGGCTCCCGAATTGTGGGGGACAAAAGCTCTGCAGTCTC 239
Qy 2094 GGCTGAGGGTCTCACCACACAAAGAGGGGAGCCGGTGTAGCAGAGGCTGAAGAGGGTG 2153
Db 240 GGCTGAGGGTCTCACCACACAAAGAGGGGAGCC----- 274
Qy 2154 GGAAGCAGGGGAGCTGTGCGTGTGTCGGGGGGGTGGAACCTTAGCGGACCTCGGAGGA 2213
Db 275 ----- 274
Qy 2214 GGCTCCCGCGCGGAACCTGCGCGACCTCCCTCCCGGCTTGCTTGCAGGCGCGCGCC 2273
Db 275 -----GGCCGCGCGCC 284
Qy 2274 CACCGCCCGCGGCATGGGGGTGAAGCCCGCCAGAGACGCTGTTCGCGTGCCTGCCA 2333
|||||

Db 285 CACGCCCGCGCCATGGGGCTGAAGCGCGCCAGAAAGACGCTGTCCCGCTGGCTCCA 344
Qy 2334 TCGACGACGTGGTGGCTGCTTTCTGCTCCGAGCTGGCGGAGAGAGCGGACCTGGTC 2393
Db 345 TCGACGACGTGGTGGCTGCTTTCTGCTCCGAGCTGGCGGAGAGAGCGGACCTGGTC 404
Qy 2394 TCCTTCTTGGTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTA 2453
Db 405 TCCTTCTTGGTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTA 464
Qy 2454 CCAAGTTCGGAGCTCACCTTCCAGCCAGCCCGCCCGCCGCGCTGGCGGCTCA 2513
Db 465 CCAAGTTCGGAGCTCACCTTCCAGCCAGCCCGCCCGCCGCGCTGGC-GCCTCA 523
Qy 2514 CCTACTTTCCGTCGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCTTCAACCGCC 2573
Db 524 CCTACTTTCCGTCGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCTTCAACCGCC 583
Qy 2574 AGATCCGAGCGCGCTGCACCTGTCCCTCTATCTCGAGAGGGGGTGTCTCCAGCGGTG 2633
Db 584 AGATCCGAGCGCGCTGCACCTGTCCCTCTATCTCGAGAGGGGGTGTCTCCAGCGGTG 643
Qy 2634 AGCTGGTGAAGAGGTCTCGATGTCTATATGGAACAGCTCAGCGCTCTCTACTTCAAG 2693
Db 644 AGCTGGTGAAGAGGTCTCGATGTCTATATGGAACAGCTCAGCGCTCTCTACTTCAAG 703
Qy 2694 ATCGGGCCCATCCAGTCCCTCTTTCAGCTTCATCAGG 2733
Db 704 ATCGGGCCCATCCAGTCCCTCTTTCAGCTTCATCAGG 743
RESULT 7
AW950638 641 bp mRNA EST 01-JUN-2000
LOCUS EST382708 MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
ACCESSION AW950638
VERSION AW950638.1 GI:8140296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 17
Seq primer: Reverse.
FEATURES
source
1..641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGA"
/note="Vector: pBluescriptSKm"
BASE COUNT 152 a 214 c 156 g 119 t
ORIGIN
Query Match 6.2%; Score 566.8; DB 122; Length 641;
Best Local Similarity 96.0%; Pred. No. 1.4e-42;
Matches 603; Conservative 0; Mismatches 22; Indels 3; Gaps 2;
Qy 7765 GACCCCGCGGAGGCTCTCTGGGACTGTGCTGGCAGACCGCGGCTTGAAGTGGCA 7824
|||||

```

Db 16 GACCCCGCCGGAAGCCTCCTGGGACTGTGCTGGCAGACAGCCCGAGGCGCTGAAGGTGGCA 75
QY 7825 GCACGGCTCAGTGGCAGCACCAGCAGCATCACACCGCGGAGGTGCAGTGCACATT 7884
Db 76 GCACGGCTCAGTGGCAGCACCAGCAGCATCACACCGCGGAGGTGCAGTGCACATT 135
QY 7885 TCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTGGTGCCACCAAGATCAACTCGA 7944
Db 136 TCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTGGTGCCACCAAGATCAACTCGA 195
QY 7945 GCGCCATCAGTGCACACTCAGCGCAGCAGTGCACAGTGCAGATGAAGAGCAGAAGTGT 8004
Db 196 GCGCCATCAGTGCACACTCAGCGCAGCAGTGCACAGTGCAGATGAAGAGCAGAAGTGT 255
QY 8005 CCACCCCTAGTGCACACTGCTCTTCTTCCCAAGCGCAGCGCAAGGCGCTCTCAACTA 8064
Db 256 CCACCCCTAGTGCACACTGCTCTTCTTCCCAAGCGCAGCGCGCAAGGCGCTCTCAACTA 315
QY 8065 CTGGGAGCTTCGGACCGCTTGTGGGAGCCAGGCTCCG-CCTTAGTCCCGCCCAACTCTGAG 8123
Db 316 CTGGGAGCTTCGGACCGCTTGTGGGAGCCAGGCTCCGCGCTTAGTCCCGCCCAACTCTGAG 375
QY 8124 CCCATGTTCTGCGCCCGCCAGCCCAAGGGGACAGGCGCTACCTCTACCCAAACCTAGGTTT 8183
Db 376 CCCATGTTCTGCGCCCGCCAGCCCAAGGGGACAGGCGCTACCTCTACCCAAACCTAGGTTT 435
QY 8184 CCGGTCCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCCAGCTCAAGACCCAGGCT 8243
Db 436 CCGGTCCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCCAGCTCAAAACCCAGGCT 495
QY 8244 CTGCCCCAGTGGTATAGATATAGTCTCTTCTCCAGATCCCGCGGCCCAATGGAAC 8303
Db 496 TTGCCCCAAGTCTGATAGATATAGTCTTCTTCCAGATCCCGCGGCCCAATGGAAC 555
QY 8304 CTCAGCTGGGTCTTAATACAGTCTTTAAGGCGCCAGCGCTAGAACCCCAACTCTT 8363
Db 556 CTCAGCTGGGTCTTAATACAGTCTTTAAGGCGCCCA-CCCCATAAAGCCCAACTTCT 613
QY 8364 COTCGGAACCGCTCACCTAGAGCCAGAC 8391
Db 614 TCTTGGAACCGGTACCTAAAGTCGGAC 641

RESULT 8
LOCUS BE267618 580 bp mRNA EST 13-JUL-2000
DEFINITION 601124837Fl NIH_MGC_8 Homo sapiens cDNA clone IMAGE:2989521 5',
mRNA sequence.
ACCESSION BE267618
VERSION BE267618.1 GI:9141211
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LUCM80 row: i column: 10
High quality sequence stop: 579.
Location/Qualifiers
1. 580
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```

```

/clone="IMAGE:2989521"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="organ: lymph; vector: pOTB7; site_1: XhoI; site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).".
BASE COUNT 128 a 197 c 174 g 81 t
ORIGIN

Query Match 6.1%; Score 560.4; DB 165; Length 580;
Best Local Similarity 99.7%; Pred. No. 5.5e-42;
Matches 572; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7640 GGGCGAGGAAGCCCGGAGGCGCGGCGGCGGCGGCGGAGTCCAGCCAGAGA 7699
Db 1 GGGCGAGGAAGCCCGGAGGCGCGGCGGCGGCGGCGGAGTCCAGCCAGAGA 60
QY 7700 GCGCCCGCGCCCAAGAACGAGCAGTGGACAAAGGCGCTGGGCGGCGGAGTCCAGCCAGAGA 7759
Db 61 GCGCCCGCGCCCAAGAACGAGCAGTGGACAAAGGCGCTGGGCGGCGGAGTCCAGCCAGAGA 120
QY 7760 GTCAGGACCCCGGAGGCGCTTCTGGGAGTGTCTGCTGGCAGCGGCGGCGGAGTCCAGG 7819
Db 121 GTCAGGACCCCGGAGGCGCTTCTGGGAGTGTCTGCTGGCAGCGGCGGAGTCCAGG 180
QY 7820 TGGCAGCAGCGCTCAGGTGCCAGCACCGCGCAGCATCACCGCGGAGGTCAGTGTCT 7879
Db 181 TGGCAGCAGCGCTCAGGTGCCAGCACCGCGCAGCATCACCGCGGAGGTCAGTGTCT 240
QY 7880 CACTTTCAGAGTGAAGAGATGAAGGCGCATGAAGGAGTGTGTGGCGGCGGAGTCCAGTCA 7939
Db 241 CACTTTCAGAGTGAAGAGATGAAGGCGCATGAAGGAGTGTGTGGCGGCGGAGTCCAGTCA 300
QY 7940 CTCGAGCGCCATCAAGTGCACGCGCAGTGCAGAGTGCAGATGAAGAGCAGAA 7999
Db 301 CTCGAGCGCCATCAAGTGCACGCGCAGTGCAGAGTGCAGATGAAGAGCAGAA 360
QY 8000 AGTGTCACCCCTAGTGTACACTGTCTTCTTCTCAAGCGGCGGCGGAGGCTCTG 8059
Db 361 AGTGTCACCCCTAGTGTACACTGTCTTCTTCTCAAGCGGCGGCGGAGGCTCTG 420
QY 8060 AACTACTGGGAGTTCGGACCGCTTGTGGGACCCAGGCTCGG-CCTTAGTCCCGCAACT 8118
Db 421 AACTACTGGGAGTTCGGACCGCTTGTGGGACCCAGGCTCGGCTTGTGGGAGTTC 480
QY 8119 CTGAGCCCATGTCTGCGCCCGCCAGCGGCGGCGGAGGCGGAGTCCAGTCCAGTCA 8178
Db 481 CTGAGCCCATGTCTGCGCCCGCCAGCGGCGGCGGAGGCGGAGTCCAGTCCAGTCA 540
QY 8179 GGTTCCTCCCGTCCCGAGTACAGTCTGTATCAACC 8212
Db 541 GGTTCCTCCCGTCCCGAGTACAGTCTGTATCAACC 574

RESULT 9
LOCUS AW968425 574 bp mRNA EST 01-JUN-2000
DEFINITION EST380501 MAGE resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
ACCESSION AW968425
VERSION AW968425.1 GI:8158266
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 574)

```

AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@igr.org
Plate: 253
Seq primer: Forward.

FEATURES
source
1..574
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4214272"
/note="Vector: pBluescriptSKm"
BASE COUNT 133 a 205 c 117 g 119 t
ORIGIN

Query Match 5.8%; Score 532; DB 122; Length 574;
Best Local Similarity 98.6%; Pred. No. 1.9e-39;
Matches 568; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 8217 ATTTTCTCAGCTCAGAACCCAGGGCTGCCCCAGTCGTTAGAAATATAGTCTCTTCTC 8276
Db 1 ATTTTCTCAGCTCAGAACCCAGGGCTGCCCCAGTCGTTAGAAATATAGTCTCTTCTC 60
QY 8277 CCAGAAATCCAGCGCGGCAATGAAACCTCACGCTGGTGCTTAATACCAAGTCTTTAAAG 8336
Db 61 CCAGAAATCCAGCGCGGCAATGAAACCTCACGCTGGTGCTTAATACCAAGTCTTTAAAG 120
QY 8337 GCCAGCCCTTAGAAACCAACCAAGCTCTCTCGGAACCGCTCACTAGAGCCAGACCAACG 8396
Db 121 GCCAGCCCTTAGAAACCAACCAAGCTCTCGCGGAACCGCTCACTAGAGCCAGACCAACG 180
QY 8397 TTACTCAGGGCTCTCCAGCTGTAGGAGCTGAGGTTTCACCCTTAACCCAGGGAGCA 8456
Db 181 TTACTCAGGGCTCTCCAGCTGTAGGAGCTGAGGTTTCACCCTTAACCCAA-GGAGCA 239
QY 8457 CAGGTCCACCTCCAGCCGGGAGCTAGGACCACTCAGCCCTCCTAGGAGTATATTCCG 8516
Db 240 CAGGTCCACCTCCAGCC-GGAGAGCTAGGACCACTCAGCCCTCCTAGGAGTATATTCCG 298
QY 8517 CACTTCAGAAATCCATATCTTGCGAATCCAAAGCTCCCTGCCCCAAATAAAGTTCAGTCTG 8576
Db 299 CACTTCAGAAATCCATATCTTGCGAATCCAAAGCTCCCTGCCCCAAATAAAGTTCAGTCTG 358
QY 8577 CTTCCAGAAATTTGGAATCCTAGTTTCCTCTCTCTGATATCCCGAGTCTGGGACACAAA 8636
Db 359 C-TCCAGAAATTTGGAATCCTAGTTTCCTCTCTCTGATATCCCGAGTCTGGGACACAAA 417
QY 8637 CTCGCCCCCAGCTATGAGATCTTGAGCCCGCCCTCTTCCTGAGCAAACTGGCCCCG 8696
Db 418 CTCGCCCCCAGCTATGAGATCTTGAGCCCGCCCTCTTCCTGAGCAAACTGGCCCCG 477
QY 8697 GATCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTC 8756
Db 478 GATCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTC 537
QY 8757 GGAGCATCCGGAGGAATCTCAGAGGGTTAGGA 8792
Db 538 GGAGCATCCGGAGGAATCTGCCAAGGGTTAGAA 573

RESULT 10
BF526328
LOCUS 799 bp mRNA
DEFINITION 602071372F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4214272
EST 11-DEC-2000

5', mRNA sequence.
BF526328
BF526328.1 GI:11613679
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabp@remail.nih.gov
Tissue procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9787 row: p column: 17
High quality sequence start: 3
High quality sequence stop: 644.
Location/Qualifiers
1..799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4214272"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.57 Kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 185 a 243 c 266 g 104 t
ORIGIN

Query Match 5.3%; Score 489.6; DB 150; Length 799;
Best Local Similarity 87.3%; Pred. No. 9.9e-36;
Matches 559; Conservative 0; Mismatches 79; Indels 2; Gaps 2;
QY 7499 GGATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGGCGGCAACCTCTCTCCTCCTT 7558
Db 83 GGAGGGCAGTCCCAACGCTGTGTGCACGTGGGTGGGCCACCTTTCTGTGCAGTCCCT 142
QY 7559 GCTCTCCCCACTGGCCAGGTGGCGAGAGGTGCGCATAGTGAGCCGAGAGGCCGAGGC 7618
Db 143 AGCCCGTTTGGAGGACAGGTGGCGAGAGGTGCGCATAGTGAGCCGAGAGGCCGAGGC 202
QY 7619 GGCCGAGCCGAGAGCCGTGGGGCGAGGAAGCCCGGGAAGGCCGGCGGGGCCACG 7678
Db 203 GGCCGAGCCGAGAGCCGTGGGGCGAGGAAGCCCGGGAAGGCCGGCGGGGCCACG 262
QY 7679 GCGGAGTCCAAAGCCAGAGAGCCCGCCGCCCAAGCCAGCCTGCACAGGCGCT 7738
Db 263 GCGGAGTCCAAAGCCAGAGAGCCCGCCGCCCAAGCCAGCCTGCACAGGCGCT 322
QY 7739 GGGCAGCCGCGCAGGTGTCAGTGTCAGGACCCCGCCGGAAGCCTCTCTGGGACTGTGCTGG 7798
Db 323 GGGCAGCCGCGCAGGTGTCAGTGTCAGGACCCCGCCGGAAGCCTCTCTGGGACTGTGCTGG 382
QY 7799 CACAGCCCGAGGCCCTGAAAGGTGGCAGCAGCGGTTCAGTGCACAGCCCGCAGCATCACC 7858
Db 383 CACAGCCCGAGGCCCTGAAAGGTGGCAGCAGCGGTTCAGTGCACAGCCCGCAGCATCACC 442
QY 7859 ACCGCCGAGGGTCCAGTGTCTACTTTCCAGAGTGAAGAGTGAAGGCGAGCT 7918
Db 443 ACCGCCGAGGGTCCAGTGTCTACTTTCCAGAGTGAAGAGTGAAGGCGAGCT 502
QY 7919 GCTGGTGGCCCAAGATCAACTCGAGCGCCATCAAGCTGCAAGCTCAGCGCACAGTCGCA 7978

```

|||||
Db 503 GCTGTGGCCACCAAGATCAAGTCGAGGCCATCAAGTCGAACTCAGCGGACAGTGCGCA 562
|||||
Qy 7979 AGTCAGATGAAGAG-CAGAAAGTCCACCCCTAGTACTACACTCTGTCTTCTCTCA 8037
|||||
Db 563 AGTCAGATGACGAAAGGTCAGAAAGTGTCACCCCTAGTACTACACTCTGATCAATC 622
|||||
Qy 8038 AGCGCAGCGCAAGGCTCTGAACACTCTGGGAGCTTCGGACCGCTTGTGGGAGCCAGG 8097
|||||
Db 623 AAGCGGAGCAAGGGCTCTGACTACGAGGACTTCGGACCGCTTGTGGGGA-CCAGG 681
|||||
Qy 8098 CTCGCCCTTAGTCCGCCAATCTGAGCCCATGTTCTGCCC 8137
|||||
Db 682 ATCCGCATAGTCCCGCAATCTGAGCCAAAGTCGAGGCC 721
|||||

RESULT 11
BF803247/c
LOCUS IL5-C10149-011100-224-g09 C10149 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF803247
ACCESSION BF803247.1 GI:12132236
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-C10149-
011100-224-g09&t3=2000-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 398.
FEATURES
Location/Qualifiers
source
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="C10149"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 107 a 107 c 189 g 118 t 2 others
ORIGIN

Query Match 5.3%; Score 484; DB 169; Length 523;
Best Local Similarity 99.08; Pred. No. 4e-35; Indels 3; Gaps 3;
Matches 517; Conservative 0; Mismatches 2;

```

```

Qy 8054 CCTCTGAACACTCTGGGACTTCGGACCGCTTGTGGGACCCAGGCTCGG-CCTTAGTCCC 8112
|||||
Db 523 CCTCTGAACACTCTGGGACTTCGGACCGCTTGTGGGACCCAGGCTCGGCTTATCTCCN 464
|||||
Qy 8113 CCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGAGGAGGCTCCTCTACCCAA 8172
|||||
Db 463 CCAACTCTGAGCCCATGTTCTGCCNCCAGCCCAAGGGAGGAGGCTCCTCCTCTACCCAA 404
|||||
Qy 8173 ACCCTAGTCTCCCGGTCGGAGTACAGTCTATCAAAACCACGATTTTCTCCAGCTCAG 8232
|||||
Db 403 ACCCTAGTCTCCCGGTCGGAGTACAGTCTATCAAAACCACGATTTTCTCCAGCTCAG 344
|||||
Qy 8233 AACCCAGGGCTCTGCCCGGCTCTAGTAATAGTCTCTTCTCCAGAAATCCAGCGG 8292
|||||
Db 343 AACCCAGGGCTCTGCCCGGCTCTAGTAATAGTCTCTTCTCCAGAAATCCAGCGG 284
|||||
Qy 8293 CCAATGGAACCTCAGCTGGGTCTTAATACAGTCTTTTAAAGGCCAGGCCCTTAGAAA 8352
|||||
Db 283 CCAATGGAACCTCAGCTGGGTCTTAATACAGTCTTTTAAAGGCCAGGCCCTTAGAAA 224
|||||
Qy 8353 CCCAAGCTCTCTCGGAACCGCTCACCTAGACGACAGCAACGTTACTCAGGGCTCCTC 8412
|||||
Db 223 CCCAAGCTCTCTCGGAACCGCTCACCTAGACGACAGCAACGTTACTCAGGGCTCCTC 164
|||||
Qy 8413 CCAGCTGTAGGAGTCTGAGGTTTCCACCTTAAACCAAGGGAGACAGGTCCAGTCCAG 8472
|||||
Db 163 CCAGCTGTAGGAGTCTGAGGTTTCCACCTTAAACCAAGGGAGACAGGTCCAGTCCAG 105
|||||
Qy 8473 CCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACCTTCAGAAATTCAT 8532
|||||
Db 104 CCC-GGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACCTTCAGAAATTCAT 46
|||||
Qy 8533 ATCTTGCGAATCCAAGTCTCCCTGCCCAATAACTTTCAGTCC 8574
|||||
Db 45 ATCTTGCGAATCCAAGTCTCCCTGCCCAATAACTTTCAGTCC 4
|||||

RESULT 12
AL521543
LOCUS AL521543 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB003YA04 5
DEFINITION prime, mRNA sequence.
ACCESSION AL521543
VERSION AL521543.1 GI:12785036
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 876)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL Contact: Genoscope
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
source
1..876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DB003YA04"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center

```


cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaide, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1019 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 486.

FEATURES

source
Location/Qualifiers
1. 525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1292871"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCACTCAAGTGGAGCGGGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaide."

BASE COUNT
ORIGIN

Query Match 5.1%; Score 466.8; DB 11; Length 525;
Best Local Similarity 98.7%; Pred. No. 1.4e-33;
Matches 523; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 8347 TAGAACCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 8406
Db 525 TAGAACCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 466

Qy 8407 CTCCTCCAGCTTGTAGAGCTGAGTTTACCCCTTACCCCAAGGAGCAGCTCCAC 8466
Db 465 CTCCTCCAGCTTGTAGAGCTGA-GTTTACCCCTTACCCCAAGGAGCAGCTCCAC 408

Qy 8467 CTCAGCCCGGGAGCCTAGGACCACTCAGCCCTTAGAGTATATTTCCGCACTTCAGAA 8526
Db 407 CTCAGCCCGGGAGCCTAGGACCACTCAGCCCTTAGAGTATATTTCCGCACTTCAGAA 349

Qy 8527 TTCCATATCTTTGCGAATCCAGCTCCCTGCCCCCAAAATTAATTCAGTCTGCTTCAGAA 8586
Db 348 TTCCATATCTTTGCGAATCCAGCTCCCTGCCCCCAAAATTAATTCAGTCTGCTTCAGAA 290

Qy 8587 TTGGAATCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAACTCCGCCCC 8646
Db 289 TTGGAATCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAACTCCGCCCC 230

Qy 8647 AGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAAACTGGCCCGGATCAGAC 8706
Db 229 AGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAAACTGGCCCGGATCAGAC 170

Qy 8707 GGACCTCCCTTCGACCTCTGCGAACCCTCCAGAGTCCAGCCCATCTCGGAGATCC 8766
Db 169 GGACCTCCCTTCGACCTCTGCGAACCCTCCAGAGTCCAGCCCATCTCGGAGATCC 110

Qy 8767 GGAGAAATCTGCAGAGGGTGTAGAGTGGGTGACAGAGCTGATCTCTCTGTTTG 8826
Db 109 GGAGAAATCTGCAGAGGGTGTAGAGTGGGTGACAGAGCTGATCTCTCTGTTTG 51

Qy 8827 TACATAGATTTATTTTTCAGTTTCAAGAAGATGAATACATTTTGTATAA 8876
Db 50 TACATAGATTTATTTTTCAGTTTCAAGAAGATGAATACATTTTGTATAA 1

RESULT 15
AI333944/c

LOCUS
DEFINITION

AI333944 525 bp mRNA EST 13-FEB-1999
qq27b09.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:193721
3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AI333944
AI333944.1 GI:4070503
EST.
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 525)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 990 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 370.

FEATURES

source

Location/Qualifiers
1. 525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:193721"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT
ORIGIN

Query Match 5.0%; Score 461.8; DB 19; Length 525;
Best Local Similarity 97.0%; Pred. No. 3.8e-33;
Matches 513; Conservative 0; Mismatches 12; Indels 4; Gaps 4;

Qy 8346 CTAGAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 8405
Db 525 CTAGAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 466

Qy 8406 GCTCTCTCCAGCTTGTAGAGCTGAGTTTACCCCTTAAACCCAGGAGCAGCTCCCA 8465
Db 465 GGTCTCTCCAGCTTGTAGAGCTGAGTTTACCCCTTAAACCCAA-GGAGCAGAGTCCCA 407

Qy 8466 CCTCAGCCCGGGAGCCTTAGAGCAGCTCAGCCCTTAGAGTATATTTCCGACATTCAGA 8525
Db 406 CCTCAGCCCGGGAGCCTTAGAGCAGCTCAGCCCTTAGAGTATATTTCCGACATTCAGA 348

Qy 8526 ATTCATATCTTGGGAATCCCAAGCTCCCTGCCCCCAAAATTAATTCAGTCTGCTTCCAGAA 8585
Db 347 ATTCATATCTTGGGAATCCCAAGCTCCCTGCCCCCAAAATTAATTCAGTCTGCTTCCAGAA 289

Qy 8586 TTTGGAATCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAACTCCGCCCC 8645
Db 288 TTTGGAATCTAGTTTCTCTCTCGTATCCCGAGTCTGGGACACAAACTCCGCCCC 229

Qy 8646 CAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTGACGAAACTGGCCCGGATCAGAGC 8705


```
|||||
Db 228 CAGCCTATGAGCATCTGAGCCCGCCCTCTCTGACGAAACTGGCCCGGATCAGAGC 169
QY 8706 AGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGGAGCATCC 8765
Db 168 AGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGGAGCATCC 109
QY 8766 CGGAGGAAATCTGCAGAGGGGTTAGGAGTGGGTGACAAGAGCCTGATCTCTCTGTTTT 8825
Db 108 CGGAGGAAATCTGCAGA-GGGTTAGGAGTGGGTGACAAGAGCCTGATCTCTCTGTTTT 50
QY 8826 GTACATAGATTTATTTTCAGTTTCCAGAAAGATGAATACATTTTGTTA 8874
Db 49 GTACATAGATTTATTTTCAGTTTCCAGAAAGATGAATACATTTTGTTA 1
```

Search completed: October 14, 2001, 03:46:30
Job time: 28096 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 08:17:47 ; Search time 2206.06 Seconds
(without alignments)
9952.996 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGCTGTGAACCTCTGCC.....AGCCACGGCGCGCGCCGCC 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2230464 seqs, 119590913 residues

Total number of hits satisfying chosen parameters: 4460928

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.1*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.2*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	9179	100.0	9180	5	US-09-380-337-3
2	9179	100.0	9180	7	US-09-880-107-3461
3	9040.4	98.5	11190	5	US-09-948-941-529
4	9040.4	98.5	11205	5	US-09-948-941-649
5	8067.2	87.9	8185	6	US-09-760-475-3945
6	1425.2	15.5	2770	1	PCT-US01-14827-388
7	1365.2	14.9	2121	1	PCT-US01-14827-5297
8	1312.2	14.3	2772	5	US-09-948-941-5
9	1312.2	14.3	2772	5	US-09-380-337-1
10	1256.4	13.7	2767	5	US-09-948-941-125
11	1252.8	13.6	2168	6	US-09-758-466-258
12	1231.4	13.4	3534	8	US-60-278-561-1743
13	698.6	7.6	3534	8	US-60-278-561-1743
14	639	7.0	639	7	US-09-864-761-22111
15	600.6	6.5	601	5	US-09-948-941-1028
16	600.6	6.5	601	5	US-09-948-941-6056
17	589.6	6.4	601	5	US-09-948-941-1037
18	589.6	6.4	601	5	US-09-948-941-6055
19	540.6	5.9	11557	1	PCT-US01-08656-3267
20	495	5.4	15957	6	US-09-891-498-3
21	473.4	5.2	519	6	US-09-878-178-1236
22	452	4.9	464	7	US-09-864-761-5337
23	414	4.5	15957	6	US-09-891-498-3
24	412.2	4.5	4227	1	PCT-US01-14827-6273
25	400.6	4.4	599	8	US-60-253-651-22088

ALIGNMENTS

RESULT 1
US-09-380-337-3
; Sequence 3, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Settara C.
; Guru, Siradanahalli C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Debelenko, Larisa V.
; Lubensky, Irina A.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated With Multiple Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,337
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-3151000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

Sequence 22376, A
Sequence 892, App
Sequence 8463, App
Sequence 216, App
Sequence 6906, App
Sequence 442, App
Sequence 564, App
Sequence 505, App
Sequence 8775, App
Sequence 8703, App
Sequence 34623, A
Sequence 696, App
Sequence 2565, App
Sequence 39315, A
Sequence 695, App
Sequence 607, App
Sequence 2104, App
Sequence 530, App
Sequence 475, App
Sequence 5725, App

LENGTH: 9180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: -
LOCATION: 1..9180
OTHER INFORMATION: /note= "genomic sequence for MEN1 gene"

FEATURE:

NAME/KEY: exon

LOCATION: 1680..1766

FEATURE:

NAME/KEY: intron

LOCATION: 1767..2264

FEATURE:

NAME/KEY: exon

LOCATION: 2265..2732

FEATURE:

NAME/KEY: intron

LOCATION: 2733..4296

FEATURE:

NAME/KEY: exon

LOCATION: 4297..4505

FEATURE:

NAME/KEY: intron

LOCATION: 4506..4715

FEATURE:

NAME/KEY: exon

LOCATION: 4716..4844

FEATURE:

NAME/KEY: intron

LOCATION: 4845..5176

FEATURE:

NAME/KEY: exon

LOCATION: 5177..5217

FEATURE:

NAME/KEY: intron

LOCATION: 5218..5297

FEATURE:

NAME/KEY: exon

LOCATION: 5298..5385

FEATURE:

NAME/KEY: intron

LOCATION: 5386..6024

FEATURE:

NAME/KEY: exon

LOCATION: 6025..6161

FEATURE:

NAME/KEY: intron

LOCATION: 6162..6622

FEATURE:

NAME/KEY: exon

LOCATION: 6623..6758

FEATURE:

NAME/KEY: intron

LOCATION: 6759..7195

FEATURE:

NAME/KEY: exon

LOCATION: 7196..7360

FEATURE:

NAME/KEY: intron

LOCATION: 7361..7577

FEATURE:

NAME/KEY: exon

LOCATION: 7578..8881

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-380-337-3

Query Match

Best Local Similarity

Matches 9180; Conservative

100.0%; Score 9179; DB 5; Length 9180;

100.0%; Pred. No. 1e-171;

0; Mismatches 0; Indels 0; Gaps 0;

1 CTGCTTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAAGTGTGTAA 60
1 CTGCTTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAAGTGTGTAA 60
61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATATCTGTCAGGTATTTG 120
61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATATCTGTCAGGTATTTG 120
121 TGGAAATCCTCAACCTGTTTTGCCAGATGTTTTCTCATGATTAGAGGAGGTATAA 180
121 TGGAAATCCTCAACCTGTTTTGCCAGATGTTTTCTCATGATTAGAGGAGGTATAA 180
181 ATTTTGGAGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
181 ATTTTGGAGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
241 ATTTTGGAGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 300
241 ATTTTGGAGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 300
301 TTTTGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 360
301 TTTTGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 360
361 TTCAAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTATCCAGCACTTT 420
361 TTCAAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTATCCAGCACTTT 420
421 TGGAGCCGAGGTAGGAGGATCACTTGAGGTGAGGAGTTCGAGACCAGCCTGACCAACAT 480
421 TGGAGCCGAGGTAGGAGGATCACTTGAGGTGAGGAGTTCGAGACCAGCCTGACCAACAT 480
481 GGTGACACCCCTGCTCTACTTAAGTAAATACAAAATACAAAATACAAAATACAAAAT 540
481 GGTGACACCCCTGCTCTACTTAAGTAAATACAAAATACAAAATACAAAATACAAAAT 540
541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAAGAATCACTTGAACCTGGAGGCG 600
541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAAGAATCACTTGAACCTGGAGGCG 600
601 AGTTGAGTGAAGCCGAGATCACTTGGGAAGCTGAGGCAGAAGAATCACTTGAACCTGGAG 660
601 AGTTGAGTGAAGCCGAGATCACTTGGGAAGCTGAGGCAGAAGAATCACTTGAACCTGGAG 660
661 TGCTCAAAACAAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACA 720
661 TGCTCAAAACAAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACA 720
721 AATTTTAGCCTTTGGGAGGCTGAGCGGGCAGATCACTTGAAGGTTAGGGTTCCAGACC 780
721 AATTTTAGCCTTTGGGAGGCTGAGCGGGCAGATCACTTGAAGGTTAGGGTTCCAGACC 780
781 AGTCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATACAAAAGTTTCTGGGTGT 840
781 AGTCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATACAAAAGTTTCTGGGTGT 840
841 GGTGGCGGAGCCCTATAATCCAGCTACTTGGGAGTCTTTTAAAGAGGAACTCACTC 900
841 GGTGGCGGAGCCCTATAATCCAGCTACTTGGGAGTCTTTTAAAGAGGAACTCACTC 900
901 TGTTGCCAGGCTGGAGTGCAGTGGCAAGATTTCTGGCTCACTGAAGCCTCCGCTCCAG 960
901 TGTTGCCAGGCTGGAGTGCAGTGGCAAGATTTCTGGCTCACTGAAGCCTCCGCTCCAG 960
961 GTTCAAGGGATTCCCGCGCCTCAGCCTCCCAAGTACCTGAGGAACTCCCTGTCTGCA 1020
961 GTTCAAGGGATTCCCGCGCCTCAGCCTCCCAAGTACCTGAGGAACTCCCTGTCTGCA 1020
1021 AAAAAAAAAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATATAT 1080
1021 AAAAAAAAAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATATAT 1080

Db	2161		GGGAGCTGTGCGTGTGTCGGGCGGGTGGAACTTATAGCGACCTTGGGAGAGGCTCC	2222
Qy	2221	CGCGGAACCTGCCACACCTCCCTCCCGGCTTGCATGCAAGCGCGCCGCCACCGCC	2280	
Db	2221	CGCGGAACTGCCCGACCTCCCTCCCGGCTTGCATGCAAGCGCGCCGCCACCGCC	2280	
Qy	2281	CGCGGCATGGGGCTGAAGCGCGCCAGAAAGCGCTGTTCGGCTCGGCTCCATCGACGA	2340	
Db	2281	CGCGGCATGGGGCTGAAGCGCGCCAGAAAGCGCTGTTCGGCTCGGCTCCATCGACGA	2340	
Qy	2341	CGTGGTGGCGCTGTTCGCGAGCTGGGCGAGAGCGGACCTGGTGTCTTC	2400	
Db	2341	CGTGGTGGCGCTGTTCGCGAGCTGGGCGAGAGCGGACCTGGTGTCTTC	2400	
Qy	2401	CTTGGTGTGGGCTCGTGGAGCATTTCTGGCTGTCAACCGCGTCACTCCCTACCAAGT	2460	
Db	2401	CTTGGTGTGGGCTCGTGGAGCATTTCTGGCTGTCAACCGCGTCACTCCCTACCAAGT	2460	
Qy	2461	TCCCGAGCTCACCTCCAGCCAGCCCGCCCGCCAGCCGCTTGGCGGCTCACCTACTT	2520	
Db	2461	TCCCGAGCTCACCTCCAGCCAGCCCGCCCGCCAGCCGCTTGGCGGCTCACCTACTT	2520	
Qy	2521	TCCCGTGGCGACCTGTCTATCATCGCGCCCTCTATGCCGCTTCAACGCCAGATCCG	2580	
Db	2521	TCCCGTGGCGACCTGTCTATCATCGCGCCCTCTATGCCGCTTCAACGCCAGATCCG	2580	
Qy	2581	AGGCGCGTGCAGCTGTCCCTCTATCCTCGAAGAGGGGTGTCTCCAGCGGTGAGCTGGT	2640	
Db	2581	AGGCGCGTGCAGCTGTCCCTCTATCCTCGAAGAGGGGTGTCTCCAGCGGTGAGCTGGT	2640	
Qy	2641	GAAGAAGTCTCCGATGTCTATGGAACAGCTCAGCGCTCTACTTCAAGATCGGGC	2700	
Db	2641	GAAGAAGTCTCCGATGTCTATGGAACAGCTCAGCGCTCTACTTCAAGATCGGGC	2700	
Qy	2701	CCACATCCAGTCCCTCTCAGCTTCATCACAGGTTGGAGCCAGTAGTGGAATCTTAT	2760	
Db	2701	CCACATCCAGTCCCTCTCAGCTTCATCACAGGTTGGAGCCAGTAGTGGAATCTTAT	2760	
Qy	2761	CCATGACCACTTCTTCAAAACCCCTCAATGTTTACAAACCTCTTAAAGAACTGTAAGC	2820	
Db	2761	CCATGACCACTTCTTCAAAACCCCTCAATGTTTACAAACCTCTTAAAGAACTGTAAGC	2820	
Qy	2821	CTTGTGAGGTTGGGAGGTATTTCCTCTTTCAGTGGGAACTGAAGCCACAGAGA	2880	
Db	2821	CTTGTGAGGTTGGGAGGTATTTCCTCTTTCAGTGGGAACTGAAGCCACAGAGA	2880	
Qy	2881	GGGAAATGATGCCAAAGTACACAGCGCATGGAGGCTGGAAGTGAAGCTGATCA	2940	
Db	2881	GGGAAATGATGCCAAAGTACACAGCGCATGGAGGCTGGAAGTGAAGCTGATCA	2940	
Qy	2941	CTTGGCTCCAATCATCAACCTCACCTCTGCGCCCTCAGCACCCGCCCTGCCACTGA	3000	
Db	2941	CTTGGCTCCAATCATCAACCTCACCTCTGCGCCCTCAGCACCCGCCCTGCCACTGA	3000	
Qy	3001	ACAGCTACAGAGTCTTAAGCATGACACAGAGGGCGGACAGATTTAGGGGGCAAGA	3060	
Db	3001	ACAGCTACAGAGTCTTAAGCATGACACAGAGGGCGGACAGATTTAGGGGGCAAGA	3060	
Qy	3061	AGATGAAATTTGGGCTGCATTTGAGGCAGTTAAACAAATATATGGCTATGAAGATTTT	3120	
Db	3061	AGATGAAATTTGGGCTGCATTTGAGGCAGTTAAACAAATATATGGCTATGAAGATTTT	3120	
Qy	3121	TTTTTTTTTTTTTGACAGGGTCTCACTCTGTCCCGCCAGGCTGGAGTGCAGTGGTG	3180	
Db	3121	TTTTTTTTTTTTTTTTTGACAGGGTCTCACTCTGTCCCGCCAGGCTGGAGTGCAGTGGTG	3180	
Qy	3181	ATCATGGCTCACTGCAAGCTCAGTCTCCCTGGGCTCAGAGATCTTCAACCTCAGCCTCC	3240	
Db	3181	ATCATGGCTCACTGCAAGCTCAGTCTCCCTGGGCTCAGAGATCTTCAACCTCAGCCTCC	3240	
Qy	3241	TGAGTAGCTGAGAGTACAGGATGCAACCGTGGTGTCTAATTTTTTTGATTTTTTTTG	3300	

Db 3241 TGAGTAGAGTACAGGATGACCGGTGGTGTGTTAAATTTTGTATTTTTTTTGG 3300
Qy 3301 TAGAGATGGTGTCTACATATGTGGCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATC 3360
Db 3301 TAGAGATGGTGTCTACATATGTGGCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATC 3360
Qy 3361 TGCCCGCTCAGTCTCCCAAAATGTGGGATTTACAGGTGTGAGCCACCCCAACTGGTGCC 3420
Db 3361 TGCCCGCTCAGTCTCCCAAAATGTGGGATTTACAGGTGTGAGCCACCCCAACTGGTGCC 3420
Qy 3421 TATGAAAAATTTTTTTTTTTTTCAGACGGGCTCTACATCTGTGCGCCAGGCTGGAGTGCA 3480
Db 3421 TATGAAAAATTTTTTTTTTTTTCAGACGGGCTCTACATCTGTGCGCCAGGCTGGAGTGCA 3480
Qy 3481 GTGGTGAACCTCGGCTCAGTGCAGGCTCTGCTCTCTGCTTTTCATGCCATTTCTCTCGCT 3540
Db 3481 GTGGTGAACCTCGGCTCAGTGCAGGCTCTGCTCTCTGCTTTTCATGCCATTTCTCTCGCT 3540
Qy 3541 CCTGCCCTCAGGCTCCTGTAGTGTGGGACTACAGAGAGCCTGCCACCATGCCGTGCTAAT 3600
Db 3541 CCTGCCCTCAGGCTCCTGTAGTGTGGGACTACAGAGAGCCTGCCACCATGCCGTGCTAAT 3600
Qy 3601 TTTTGTGGATTTTGTAGTAGACGAGGTTTCACCATGTTAGCCAGGATGCTCGATCT 3660
Db 3601 TTTTGTGGATTTTGTAGTAGACGAGGTTTTCACCATGTTAGCCAGGATGCTCGATCT 3660
Qy 3661 CTTGACCTCGTGATCCGCCCGCTTGGGCTCCCAAGTGTGGGATTTACAGCGGTGAGCC 3720
Db 3661 CTTGACCTCGTGATCCGCCCGCTTGGGCTCCCAAGTGTGGGATTTACAGCGGTGAGCC 3720
Qy 3721 ACCGCACTCGTCAAAAATGTTTGAGACAGAGAGGGGCTTGACCTCAAAAAGGCTTAAGA 3780
Db 3721 ACCGCACTCGTCAAAAATGTTTGAGACAGAGAGGGGCTTGACCTCAAAAAGGCTTAAGA 3780
Qy 3781 GTCAGGGCTTCCAAAGAGCTTTGCACCAAGCCCGTGTACCTGGCAATCCCATCTGGTGT 3840
Db 3781 GTCAGGGCTTCCAAAGAGCTTTGCACCAAGCCCGTGTACCTGGCAATCCCATCTGGTGT 3840
Qy 3841 GCCATATTTGAGAAGGAATCAGAGGCTGCTCTCAGCTTTAGCAGGAAAGAGTGCAGAGAT 3900
Db 3841 GCCATATTTGAGAAGGAATCAGAGGCTGCTCTCAGCTTTAGCAGGAAAGAGTGCAGAGAT 3900
Qy 3901 AAATGAGGGTTATTTGTGGTGGTGTATAGCCAGAGAGTGTGGCCAGCGTCTGTTTTT 3960
Db 3901 AAATGAGGGTTATTTGTGGTGGTGTATAGCCAGAGAGTGTGGCCAGCGTCTGTTTTT 3960
Qy 3961 TGCCATTCCTGTTTAACTTAGTAGTGCAGTAAATGSAATCCCTAAATCCATAGAATA 4020
Db 3961 TGCCATTCCTGTTTAACTTAGTAGTGCAGTAAATGSAATCCCTAAATCCATAGAATA 4020
Qy 4021 TATAATAGAGTTGCAGAGAAAGACGAGGTAGGGCCAAAGGCTGGGTGAGTACAGGATAT 4080
Db 4021 TATAATAGAGTTGCAGAGAAAGACGAGGTAGGGCCAAAGGCTGGGTGAGTACAGGATAT 4080
Qy 4081 CCAGAAAGGTATCTTTGTGGACATAGAGGGTGTAAACAGGAGAGAGTCTTTGAACAGT 4140
Db 4081 CCAGAAAGGTATCTTTGTGGACATAGAGGGTGTAAACAGGAGAGAGTCTTTGAACAGT 4140
Qy 4141 GGGAGGGAAGGATGGAGGATAGTGGCCAGGAGATCTGAGTTGGGTGCAGGCTTGG 4200
Db 4141 GGGAGGGAAGGATGGAGGATAGTGGCCAGGAGATCTGAGTTGGGTGCAGGCTTGG 4200
Qy 4201 AAAGGAGTGGAGGAGTGTGGCCCATCACTACCTGCGCCCTTTCCCATGTTAAAGCA 4260
Db 4201 AAAGGAGTGGAGGAGTGTGGCCCATCACTACCTGCGCCCTTTCCCATGTTAAAGCA 4260
Qy 4261 CAGAGGACCTCTTTTCATTAATCTCCCTTCCACAGGCAACAAATTTGACAGCTCCGGTG 4320
Db 4261 CAGAGGACCTCTTTTCATTAATCTCCCTTCCACAGGCAACAAATTTGACAGCTCCGGTG 4320
Qy 4321 TGGCCCTTGTGTGGTGTGGGCTCCAGGCGCTCGGGTCTCCGGGATGTCACCTCGGCC 4380
Db 4321 TGGCCCTTGTGTGGTGTGGGCTCCAGGCGCTCGGGTCTCGGGATGTCACCTCGGCC 4380

Qy 4381 TGTCTGAGGATCATGCTGGGTAGTGTGGGCCCAATGGGAGCAGATATGAGGTCA 4440
Db 4381 TGTCTGAGGATCATGCTGGGTAGTGTGGGCCCAATGGGAGCAGATATGAGGTCA 4440
Qy 4441 CCTGGCAGGCAAGGGCAACGAGGACCGCAGGGGCCAGAGTCAATGCCGTCTGGGTG 4500
Db 4441 CCTGGCAGGCAAGGGCAACGAGGACCGCAGGGGCCAGAGTCAATGCCGTCTGGGTG 4500
Qy 4501 AGCGGGTATTTCTCCCTCCCGCCAGCCTTGTCCCTTTCATCTAGTAGTCCAGCCAC 4560
Db 4501 AGCGGGTATTTCTCCCTCCCGCCAGCCTTGTCCCTTTCATCTAGTAGTCCAGCCAC 4560
Qy 4561 CCAAGGGACTCCATTTCTTGGGCCACACCCCTTTCTTCCCATCACCCACATAGGAA 4620
Db 4561 CCAAGGGACTCCATTTCTTGGGCCACACCCCTTTCTTCCCATCACCCACATAGGAA 4620
Qy 4621 GGGAGACAGAGAGCCCTTTTCTGGCTGTCTATTCCTTGAAGCAGCAGAGGTGGG 4680
Db 4621 GGGAGACAGAGAGCCCTTTTCTGGCTGTCTATTCCTTGAAGCAGCAGAGGTGGG 4680
Qy 4681 CATCATGAGACATATGATCTATCCCGCCCTTAAGAGCTGGCTTACCTGAAGGATCAT 4740
Db 4681 CATCATGAGACATATGATCTATCCCGCCCTTAAGAGCTGGCTTACCTGAAGGATCAT 4740
Qy 4741 ACATCGCTGTGACCGCAAGATGGAGGTGGGCTTCATGCTGTGCCATCAACCTTTCCA 4800
Db 4741 ACATCGCTGTGACCGCAAGATGGAGGTGGGCTTCATGCTGTGCCATCAACCTTTCCA 4800
Qy 4801 TTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGAGGTGAGGGTGAACAA 4860
Db 4801 TTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGAGGTGAGGGTGAACAA 4860
Qy 4861 TGGGACGAGCTGGGCTAGGCAAGCTTGTGCTGTGGAGCCCTGGGAGGGGACCTT 4920
Db 4861 TGGGACGAGCTGGGCTAGGCAAGCTTGTGCTGTGGAGCCCTGGGAGGGGACCTT 4920
Qy 4921 TCCCTTCTGAGCTTTCAGCTTCCCTTCCGAAATGGGTAGTAAATTCCTTGGCTGGC 4980
Db 4921 TCCCTTCTGAGCTTTCAGCTTCCCTTCCGAAATGGGTAGTAAATTCCTTGGCTGGC 4980
Qy 4981 CTTTCCCAGGCTCTTGGGAGTAGAATGAGATGTGAAATTCCTTGGCTGGCTTAA 5040
Db 4981 CTTTCCCAGGCTCTTGGGAGTAGAATGAGATGTGAAATTCCTTGGCTGGCTTAA 5040
Qy 5041 GGGCTGGTCCCAGAAATTTTGGCCCTTCCACATGGTGGTGGTCCCTGTTGGTCTGACCC 5100
Db 5041 GGGCTGGTCCCAGAAATTTTGGCCCTTCCACATGGTGGTGGTCCCTGTTGGTCTGACCC 5100
Qy 5101 CCACCTCTGCCCGATAGGCTAAGGACCCGTTCTCCTCCCTGTCCGTCGCTCAATCT 5160
Db 5101 CCACCTCTGCCCGATAGGCTAAGGACCCGTTCTCCTCCCTGTCCGTCGCTCAATCT 5160
Qy 5161 CTCCTTCCGCTCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
Db 5161 CTCCTTCCGCTCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
Qy 5221 AGTAGAGGAAAGTGGCCAGGCTGGCCCTTGGAGCCGGGGGCTGGGTGGGAGCTGAA 5280
Db 5221 AGTAGAGGAAAGTGGCCAGGCTGGCCCTTGGAGCCGGGGGCTGGGTGGGAGCTGAA 5280
Qy 5281 TTATGATCTCTTCTTCTAGTACCCCATGCTTGGGAAACCTGGGAGATCTAGAGAGCTG 5340
Db 5281 TTATGATCTCTTCTTCTAGTACCCCATGCTTGGGAAACCTGGGAGATCTAGAGAGCTG 5340
Qy 5341 GAGCCACCCCTTGGCCGCGCAGACCCATCAGCTTACCACAGGCTGGGAGCTTAAG 5400
Db 5341 GAGCCACCCCTTGGCCGCGCAGACCCATCAGCTTACCACAGGCTGGGAGCTTAAG 5400
Qy 5401 GAGGCTGAGAGGAGAGCCCTTAACAGTGGGTGAGGAGGGGCTTCTTGGGAGATG 5460
Db 5401 GAGGCTGAGAGGAGAGCCCTTAACAGTGGGTGAGGAGGGGCTTCTTGGGAGATG 5460

QY 5461 AGAAGAACTTTGTGTGTGGGGGTATGCCCATCCAGTCTCACTTTGTGTCAACTGT 5520
DB 5461 AGAAGAACTTTGTGTGTGGGGGTATGCCCATCCAGTCTCACTTTGTGTCAACTGT 5520
QY 5521 GTGCAGAAATCAGTTTCAGTCAAGGGCTGTCTGAGGGGTGTCCAGGGTTCCCGAGCCTGGGAG 5580
DB 5521 GTGCAGAAATCAGTTTCAGTCAAGGGCTGTCTGAGGGGTGTCCAGGGTTCCCGAGCCTGGGAG 5580
QY 5581 TGGCAGGGCTGCATTTGTCCCTCAGCCCTGACCTGCTTTCTGCCACTCTTACTGTCTTC 5640
DB 5581 TGGCAGGGCTGCATTTGTCCCTCAGCCCTGACCTGCTTTCTGCCACTCTTACTGTCTTC 5640
QY 5641 CTGAGATATAACAGAGTCAAAATGTGTAGGAGCAGTGAAGAGGGGTGTTCACCTTGGTG 5700
DB 5641 CTGAGATATAACAGAGTCAAAATGTGTAGGAGCAGTGAAGAGGGGTGTTCACCTTGGTG 5700
QY 5701 GGTGTAGGTGGGAGGAGGCCATTTGGGCTTGGGCTTGAAGTCTTTGGTGTGATGTAGAA 5760
DB 5701 GGTGTAGGTGGGAGGAGGCCATTTGGGCTTGGGCTTGAAGTCTTTGGTGTGATGTAGAA 5760
QY 5761 GAGTGTCTGAGAAAGAAAGGGCCCTGAGCTCGGAGGCGAGGCCACCCCTGCAGTCTG 5820
DB 5761 GAGTGTCTGAGAAAGAAAGGGCCCTGAGCTCGGAGGCGAGGCCACCCCTGCAGTCTG 5820
QY 5821 CCCAGGCCCTCAGCCAGCAGTCTGTAGACCAGGAGGAGACACAGTGAAGGGCTGGC 5880
DB 5821 CCCAGGCCCTCAGCCAGCAGTCTGTAGACCAGGAGGAGACACAGTGAAGGGCTGGC 5880
QY 5881 AGCAGTGGAGGTGGAGTGGAGATGAGAGGACTCCTCGGGATCTTCTGTGTGGCCCTT 5940
DB 5881 AGCAGTGGAGGTGGAGTGGAGATGAGAGGACTCCTCGGGATCTTCTGTGTGGCCCTT 5940
QY 5941 CTGGGTGTGCCCTGTGGGCTTTGTGCCAGCAGGCGAGTGGGCTGCCCTCCCTGAGG 6000
DB 5941 CTGGGTGTGCCCTGTGGGCTTTGTGCCAGCAGGCGAGTGGGCTGCCCTCCCTGAGG 6000
QY 6001 ATCCTCTGCCCTCACTCCATCCAGGGCATTTGCCCTCAGCAGGCGAGTGGGCTGCCCTCCCTGAGG 6060
DB 6001 ATCCTCTGCCCTCACTCCATCCAGGGCATTTGCCCTCAGCAGGCGAGTGGGCTGCCCTCCCTGAGG 6060
QY 6061 CACATCTACCCCTACATGTACCTGGCTTGGCTTACCACTGTGCGAACCCGAATGTGCGGGAA 6120
DB 6061 CACATCTACCCCTACATGTACCTGGCTTGGCTTACCACTGTGCGAACCCGAATGTGCGGGAA 6120
QY 6121 GCCCTGAGGCTGGGCGGACACGGCCACTGTATCATCAGGAGTCAAGTCCCTACTAG 6180
DB 6121 GCCCTGAGGCTGGGCGGACACGGCCACTGTATCATCAGGAGTCAAGTCCCTACTAG 6180
QY 6181 GGCCTGAGGCTGTCTTCTTCCCTCCCATCAGTTTCCAAACCCCTCGTCCAGGACTG 6240
DB 6181 GGCCTGAGGCTGTCTTCTTCCCTCCCATCAGTTTCCAAACCCCTCGTCCAGGACTG 6240
QY 6241 AGGCCTGGCTCCAGGCCCATCCCTTTCATCCAGTCCGTAGGAGCAAGGCCACCAT 6300
DB 6241 AGGCCTGGCTCCAGGCCCATCCCTTTCATCCAGTCCGTAGGAGCAAGGCCACCAT 6300
QY 6301 TACCAGGAGTGGGACCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCTC 6360
DB 6301 TACCAGGAGTGGGACCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCTC 6360
QY 6361 CTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAAATGTTTAAACCAACCATATCCAGC 6420
DB 6361 CTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAAATGTTTAAACCAACCATATCCAGC 6420
QY 6421 AGTGGACTTCCACCTCGGCCCATGCCCTTCTCATTTCTTCTTCTCTCTGGG 6480
DB 6421 AGTGGACTTCCACCTCGGCCCATGCCCTTCTCATTTCTTCTTCTCTCTGGG 6480
QY 6481 CTGACCCAGACAGCATATTTTGCAGTGAGGACCCCACTACTCCCCAGCCCTGGGG 6540
DB 6481 CTGACCCAGACAGCATATTTTGCAGTGAGGACCCCACTACTCCCCAGCCCTGGGG 6540
QY 6541 CTCATCCCCCGCAGGTTCCTTGGGGCTACCCCGGATGGTGAGAGCCCTTCAGAGCCCTAC 6600

DB 6541 CTCATCCCCCGCAGGTTCCTTGGGGCTACCCCGATGGTGAGAGCCCTTCAGAGCCCTAC 6600
QY 6601 AGAGACCCCACTGTCTCTACAGACTACAACTACTGCCGGGAAGAGAGAGATCTACAGG 6660
DB 6601 AGAGACCCCACTGTCTCTACAGACTACAACTACTGCCGGGAAGAGAGAGATCTACAGG 6660
QY 6661 AGTTCTTTGAAGTAGCAATGATGATCTCCCAACCTGCTCAAGGAGGACGCCAGCTTGC 6720
DB 6661 AGTTCTTTGAAGTAGCAATGATGATCTCCCAACCTGCTCAAGGAGGACGCCAGCTTGC 6720
QY 6721 TGGAGGCGGCGAGGAGCGCGGGGAGCAAAAGCCAGGTGAAAGGCTTGGAGCTCCAGCC 6780
DB 6721 TGGAGGCGGCGAGGAGCGCGGGGAGCAAAAGCCAGGTGAAAGGCTTGGAGCTCCAGCC 6780
QY 6781 TGTGTCCAGCCTCCCACTTGGAGAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA 6840
DB 6781 TGTGTCCAGCCTCCCACTTGGAGAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA 6840
QY 6841 CGGATTAGGATGGCAGGAGAGGTGGCCCTGAGCAGACAGCTATCTTCCCTTTTGTCT 6900
DB 6841 CGGATTAGGATGGCAGGAGAGGTGGCCCTGAGCAGACAGCTATCTTCCCTTTTGTCT 6900
QY 6901 ATAACTCAGGTCTCTGGGCCACGTTTGGACGGGACTGAAGGTATTTAGAGGTTTCTACCC 6960
DB 6901 ATAACTCAGGTCTCTGGGCCACGTTTGGACGGGACTGAAGGTATTTAGAGGTTTCTACCC 6960
QY 6961 TGTGCTTCAGTTTCATGCCCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGGATGG 7020
DB 6961 TGTGCTTCAGTTTCATGCCCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGGATGG 7020
QY 7021 TACCTCTGGCTATGATTTGGCTTTTATAAAGAAAGAGGTCTTAAGAAATGTTCACCAACC 7080
DB 7021 TACCTCTGGCTATGATTTGGCTTTTATAAAGAAAGAGGTCTTAAGAAATGTTCACCAACC 7080
QY 7081 TATGCTTACCTTTTCTTGAGCCAGGGGTCTTTGGCTAGGTGGGGGCTGGGCTGTGCC 7140
DB 7081 TATGCTTACCTTTTCTTGAGCCAGGGGTCTTTGGCTAGGTGGGGGCTGGGCTGTGCC 7140
QY 7141 TCTCTAAGGGTCAGTAAGAGACTGATCTGCTGCCCTCCCTTCCCTCCCTGCTCCAGGGCAC 7200
DB 7141 TCTCTAAGGGTCAGTAAGAGACTGATCTGCTGCCCTCCCTTCCCTCCCTGCTCCAGGGCAC 7200
QY 7201 CCAGAGCAAGGTTCGCGCTCCAGGACCTGAGTGTCTCCGCCACCTGCTGCGATTCTA 7260
DB 7201 CCAGAGCAAGGTTCGCGCTCCAGGACCTGAGTGTCTCCGCCACCTGCTGCGATTCTA 7260
QY 7261 CGACGGCATCTGCAATGGGAGGAGGCGAGTCCCAACGCTGTGTGTGACGTGGGCTGGGC 7320
DB 7261 CGACGGCATCTGCAATGGGAGGAGGCGAGTCCCAACGCTGTGTGTGACGTGGGCTGGGC 7320
QY 7321 CACCTTTCTGTGAGTCCCTTAGGCCGTTTGAAGGAGAGGTGAGGAGACCTGACACAGA 7380
DB 7321 CACCTTTCTGTGAGTCCCTTAGGCCGTTTGAAGGAGAGGTGAGGAGACCTGACACAGA 7380
QY 7381 GGTCTGGGCACTACAGTGTGTGACAGCAGCAGCGGCTGTGACACTTTTCTGGCCCCAGG 7440
DB 7381 GGTCTGGGCACTACAGTGTGTGACAGCAGCAGCGGCTGTGACACTTTTCTGGCCCCAGG 7440
QY 7441 GGCAGCATCTGCCCATCCCTTTCGGTGGCGATGGGACTGAGACCCCTTGGGTGGATGGG 7500
DB 7441 GGCAGCATCTGCCCATCCCTTTCGGTGGCGATGGGACTGAGACCCCTTGGGTGGATGGG 7500
QY 7501 ATGCCAGAGCAGGTCCTTGGAGTTCACGACCTTGGCCGCAACCTTCTCTACCTTGC 7560
DB 7501 ATGCCAGAGCAGGTCCTTGGAGTTCACGACCTTGGCCGCAACCTTCTCTACCTTGC 7560
QY 7561 TCTCCCCACTGGCCAGTGGCGGAGAGGTGCGCATAGTGAGCGAGAGCGCGGCGG 7620
DB 7561 TCTCCCCACTGGCCAGTGGCGGAGAGGTGCGCATAGTGAGCGAGAGCGCGGCGG 7620
QY 7621 CCGAGGCGAGGAGCTGTGGGGGAGGAGCCCGGGAAGCGCGCGGGCGGCCACCGC 7680

1201	gagcccgctgaagtcgaggtcgagtcgagccagatcgagccactgcaatlccagccttgggc	1261
1261	GAAAGAGAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTTAGCACCGTGTTAGA	1320
1261	gaaagagaagaccgtgtctcaaaacaacaacaaagcttaclctlagcacgttgtaga	1320
1321	GTATCTCGCGGCGGAAGTGGGAACAGAGTCTGCACACACAGTAGGCATCTTTATATGT	1380
1321	gtatctcgcgcggaagtgggaaacagagtcgtcacacagtaggcatactttatatgt	1380
1381	TAAACAGACACTGATACCCAGCTTAAAGCGGCTGAACACATTTTACTCTGCGCAGTGTAA	1440
1381	taacagacactgataccacagctaaagcgctgaacacatttaclctctgacagtggttaa	1440
1441	AAGTATCTGTTTTCTCATTTGTTTTATTTATTTTCTGGATCAAGCAACCTGATC	1500
1441	aagtatctgtttctcatattgtttatttaatttttctggatcaagcaactgac	1500
1501	TTTTTTCCTCATTAACCTGCGACCGACCCGTGCACGCAAAACCGCAGAGCTCGCGCAC	1560
1501	tttttctcataacttgcgacccagcctgacagcaaaacccgcagaagctcgcgacc	1560
1561	TCCCACCCGAGTCTCAGGTAGTCCCGGACCTACATTTTCCAGAAAGGCACCTCGCGGC	1620
1561	tccaccccgagctcgaggtagtgcccccgagctacatttccagaaggaacttgcgggc	1620
1621	ACGCTTCTCCCTGCTGGCCCTCAAGGAAGGGCCATCCCTGTAGTATCTCGGGAAGGAG	1680
1621	acgcttcttgccttgcgscctggaagggaaggcccaatcctctgagtatctcggaaggag	1680
1681	GTGTCGGAGCGCGGACCTAGAGATCCAGAGCCACAGCGCAGCGCCCGCGGCCGCCA	1740
1681	gtgtccgagcgcgagacctagatctccagaagccaagcgagcgcccgcccgcca	1740
1741	CTATTTCCAGGCTCTCGGGGCGAGGGTGGGCCACAGACTCCACTTCCCGCGCGGTAGTGC	1800
1741	ctatttccaggtctctcgggcgaggggtggggcccaagactccaactccgcgagtagtgc	1800
1801	GACCCTAGGGGCGGGACTTCATGTCCCAGCAGGCTCCGGGCGGGCTGGCGCGGTCCT	1860
1801	gaccttagggcgggacttcatgtcccagcaggtcccgggcgggctgcgcgcgctgctct	1860
1861	AGTGTGGGATCTAAGCGCGGAGGTGGCGAGGGGACCGAGGCCAGGACTCTCTCTGGGG	1920
1861	agtgtaggagtgaagcgcgaggtggcgaggggagccagcgagcaggaactctcttgggg	1920
1921	TTTGGGGCTTGACCTGGCTTCTTGTGGACAGACTTTACAGCCCCGGGGGCACAGT	1980
1921	tttggggcttgaccttggcttgcgtcttctggacagactttacagcccccggggcaagct	1980
1981	CGTAGAGAGGGGCGGGCGGCCATTGGGGCTCTCTATTGGGGTGTCTTGGGGCGCACCCC	2040
1981	cgtagagagggcgggcgcgccattggggctcctcatttgggtgcttggggcgacccc	2040
2041	ATCGGGTACGGGCGTCCCGGAATTGTGGGGGACAAAAGGCTCTGCGAGTCTCGGCTGAG	2100
2041	atcgggtacccggcgtcccggaaattgtggggacaaaaagcctctgcagctcggcgtgag	2100
2101	GGGTCTCACCGACAANAAGAGGGAGCCCGTGCAGCAGAGGCTGAGAGGGTGGGAAGCA	2160
2101	gggtctcacccgacaaaagaggggaagcgtgagcagagcgttgaagaggggtggggaagca	2160
2161	GGGGAGTGTGCGCTGTCTCGGGCGGGTGGAACTTACGGACCCCTGGGAGGAGCTCCC	2220
2161	ggggagctgtgcgtgtgtcgggcggggtggaaaccttagcgaccttgggaagagctccc	2220
2221	CGGCCGAACCTGCCGACCTCCCTCCCGCGCTTGCCTTGCAGGCCGCCGCCACCGCC	2280
2221	cgggcgaaacctgccgacctcctctcccgcgcttgccttgcagtcgagcgcgcccaacgccc	2280
2281	CGCGGCCATGGGCTGAAGCCGCCCAGAAAGCGTGTTCGGCTGCCTCATCTGACGA	2340
2281	cgcgcgcattgggctgaagcgcccgccagagacgctgttccgcgtgcgctcatctcagca	2340

QY 2341 CCGTGGCCCTGTTTGGTGGAGCTGGCGGAGAGAGCGGACCTGGTGTCTCTTC 2400
DB 2341 cgtggcgccctgttggctggagctggcgagagagcgagcgacctgggtgctctcttc 2400
QY 2401 CTTGGTGTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCGTATCCCTACCAAGT 2460
DB 2401 cttgggtgtggcttcgtggagcatttctggctgtcaaccggtcatccctaccaagt 2460
QY 2461 TCCCGAGCTCACTTCCAGCCCGCCGCGCCGCGCCGCGCTGCGGCGCTCACACTCT 2520
DB 2461 tcccgagctcaacttccagcccgccccgccccgccccgccccgccccgccccgccccg 2520
QY 2521 TCCCGTGGCGGACCTGTCTATCATCGCGCGCTCTATCGCGCTTACCGCGCCAGATCCG 2580
DB 2521 tcccggtggagactgtctatcatcgccgccccctctatcgccgcttcacgccccagatccg 2580
QY 2581 AGGCGCGTGGACCTGTCCCTCTATCTTCGAGAGAGGGGTGTCTCCAGCGGTGACTGGT 2640
DB 2581 aggcgcgtcgacctgtccctctatctctcgagaggggtgtctccagccgtgagctggt 2640
QY 2641 GAAGAAGTCTCCGATGTCAATGAACAGCTCAGCGCTCCTACTTCAAGGATCGGSC 2700
DB 2641 gaagaaggtctccgagtcatatgaacagcctcagcgcctcctactccaagga cggggc 2700
QY 2701 CCACATCCAGTCCCTTTCAGCTTTCATCAGGTTGGAGCGAGTAGGTGGGAATCTTTAT 2760
DB 2701 ccacatccagtcctctctcagcttcatcacaggttggagccagtagtgggaaactcttat 2760
QY 2761 CCATGACCCACTTCTCAAAACCTTCATGTTTACAGAACCTTTTAAAGACTGTAAAGC 2820
DB 2761 ccatgaccacttctcaaaacctccatggtttacagaaaccttttaagaactgaagc 2820
QY 2821 CTTGTGAGTTCGGCAGGTGTTATTTCTCTCTTTCAGTGGGAACTGAAGCCGACAGAGA 2880
DB 2821 cttgtgaagtttcggcaggtgttatctctctctctctctctctctctctctctctctct 2880
QY 2881 GGGGAAATGATATGCCAAGTACACAGCGCATGCGAGGCTGGAGGTGAAGCTGTATCA 2940
DB 2881 ggggaaatgatagccaagtacacacagcgcatgagcgaggtggagtggaagtggaagcctgatca 2940
QY 2941 CTTGGCTCCAATCATCACTCACTCTGCTCCCTCAGCAGCCGCTGCGCCTTGGCAGTGA 3000
DB 2941 cttggctccaaatcatacctctcaactctgtccccctcagccccccaccccttgccactga 3000
QY 3001 ACAGCTACAGGAGTTCTAAGCATGACACAGAGCGGCGGACAGATTAGGGGCAAGA 3060
DB 3001 acagctacaggaattctaagcatagacacagagggcgagcagattttaaggggcaaga 3060
QY 3061 AGATGAATTTGGCTGCATTTGAGGCAGTTAAACAAAATATGGCTATGAAGATTTTTTT 3120
DB 3061 agatgaatttggctgcatttgagcagtttaacaaaaataatggctatgaagattttttt 3120
QY 3121 TTTTTTTTTTTTTTCAGACAGGCTCTACTCTGCTCCCGCCAGGCTGGAGTGCAGTGGTGG 3180
DB 3121 ttttttttttttttcagacaggtctactctgtccccccagcgctggagtgagtggtgtg 3180
QY 3181 ATCATGGCTCACTGAGGCTCAGTCTCCCTGGGCTCAGAGATCTCCCACTCAGCCCTCC 3240
DB 3181 atcatggctcaactgagcctcagctcctcctgggtcagagatcctccaaacctcagcctcc 3240
QY 3241 TGAGTAGCTGAGATGACAGCATGACCGTGGTGTCTTAATTTTTTTTGTATTTTTTTTGG 3300
DB 3241 tgagtagctgagatgacagcatgacccgtggctggtgttaatttttttgtatTTTTTTT 3300
QY 3301 TAGAGATGGTGTCTACATATGTGGCCAGACTGGTCTTGAACCTCTTGGCTCAAGTGATC 3360
DB 3301 tagagatgggtgtctacatattgtggccagactgggtcttgaacctctctgggtcgaagtgtac 3360
QY 3361 TGCCCGCTCAGTCTCCCAATGTGGGATTACAGGTGTGAGCCACCGCAACTGTGTGGCC 3420
DB 3361 tgcccgctcagctctcccaatgtgggattacaggtgtgagccaccgcaactggtggcc 3420

QY 3421 TATGAAAATTTTTTTTTTTTTCAGACGGGCTCTCACTCTGTGCGCCAGGCTGGAGTGCA 3480
DB 3421 tatgaaaatTTTTTTTTTTTTTTTTCAGACGGGCTCACTCTGTGCGCCAGGCTGGAGTGCA 3480
QY 3481 GTGGTGCAATCTCGGCTCAGTCAAGCTCTGCTCCTCTCTCTTTCATGCGCATTTCTCTCGCT 3540
DB 3481 gtggtgcaatctcggctcactgcaagctctgctcctctgtttcatggcatctctcgtcct 3540
QY 3541 CTTGGCTCAGGCTCCTGAGTAGTGGGACTACAGAGAGCTGCCACCATGCGCTGGCTGAAT 3600
DB 3541 cttggctcagcctcctgagtagctgggactacagagagcctgccaccatgctctggcctaact 3600
QY 3601 TTTTTTGGATTTTTTAGTAGAGACGAGGTTTACCATGTTAGCCAGGATGTCTCTGATCT 3660
DB 3601 ttttttggatTTTTTAGTAGAGAGGTTTCAccatgtttcaaccatgttagcagagtaggtctcyalet 3660
QY 3661 CTTGACCTCTGTATCCGCGCGCTTGGGCTTCCCAAGTGTCTGGGATTTACAGGCTTGAGCC 3720
DB 3661 cttgacctgtatccgccccgcttggctcccaagtgtcgggattacagggctgaagcc 3720
QY 3721 ACCGACCTGTGTCAAAAATGTTTGACAGACAGAGGGCTTGACCTCAAAAGGCTTAAAGA 3780
DB 3721 accgacctgtgtcaaaaatgtttgagacagagaaaggcttgacctcaaaaggcttaaga 3780
QY 3781 CTCAGGGCTTCAAAAGACTTTTGACCAAGCCGCTTGACTGGCAATCCCATCTCTGCTGT 3840
DB 3781 gtcagggtctgcaaaagtcttgaccaaagcccgttgaactggcaatcccatcctctgggt 3840
QY 3841 GCCATATTGAGAGGAATCAGAGGCTGCTTCTCAGCTTTAGCAGGAAAGAGTGCAGAGAT 3900
DB 3841 gccatattgagaaggaatcagaggctgtctcagcttagcaggaaagagtgagagat 3900
QY 3901 AAATGAGGTTATTGTTGGTGGTGTATAGCACAGAGTGTGGCCAGCGCTCTGTTTTT 3960
DB 3901 aaatgagggttatgttgggtgggtgtatagccagagagtgctggccagagctctctct 3960
QY 3961 TGCCATTCCTGTTTAACTAGTAGTAAATGGAATCCCTAAATCAATCAATGAATA 4020
DB 3961 tgccattcctgttttaacctagtagtgagtaaagtgaataccctaaatcccatagaaata 4020
QY 4021 TATAATAGATTGACAGAGAAAGACGAGGTAGGGCCAAAGGCTGGGTGACGTACAGGATAT 4080
DB 4021 tataatagattgacagagaaagacgagtagggccaaaggctgggtcagagtagat 4080
QY 4081 CCAGAAAGGTATCTTGTGGACATAGAGGGTGTAAACAGGGAGAGAGTCTTTGAACAGCT 4140
DB 4081 ccagaaaggtatcttgttggacatagaggtgtaaacagggagagagctcttgaacagct 4140
QY 4141 GGGAGGGAAGGATGGAGGGATAGTGGCAGAGAGATCTGAGGTTGGGTACAGGCTTTGG 4200
DB 4141 gggagggaaggaaggagggatagtgggcagagagaaatctgaggttgggtcagaggtcttg 4200
QY 4201 AAAGGAGTGGGAGGAGTGTGGCCCATCACTACCTGCGCCCTTTCCCATCTTTAAAGA 4260
DB 4201 aaaggagtgggagggagtggtggcccatcactacacctggcccccttcccccaagttaaga 4260
QY 4261 CAGAGGACCTCTTTTCAATTACCTCCCTTCCACAGGCAACCAATTTGACAGCTCCGGTG 4320
DB 4261 cagaggacctcttcttattacctcccccttccacagggcaccaaaattggagagctccgggtg 4320
QY 4321 TGGCCTTTGCTGTGTTGGGGCTGCCAGGCGCTGGGTCTCCGGGATGTCCACCTCGGCC 4380
DB 4321 tggccttctgtgtgttgggctcagggcctgggtctccggagtggtccacctcggcc 4380
QY 4381 TGCTGTAGGATCATGCTGGGTAGTGTGTTGGGCCCAATGGGGAGCAGACAGCTGAGGTCA 4440
DB 4381 tgtctgaggaatactctgggttagtgtttggggccaaatggggagcagacagctggaggtca 4440
QY 4441 CCTGGCACGGAAGGGCAACGAGGACCGGCGGCGGACAGTCAATGCCGCTGTGGCTG 4500
DB 4441 cctggcacgggcaaggggcaacgagggcaggggccagacagtcgaatggcgggtgtggctg 4500
QY 4501 AGCGGTTATTGTTCTCCCTCCCGCCAGCCTTGTCCTCTTACTACTGTAGTACCCAGCCAC 4560

|||||
Db 4501 agcgggtattgttccctccccagcctgttcccttcaactgtagtagcccaagccac 4560
QY CCAAGGGACTCCATTCTTTGGGCCACACCCCTTCTTCCCATCACCCACATAGGAA 4620
Db ccaaggagctccattcttcttgggcccacaccccttcttcccatcaccacccacataggaa 4620
QY GGAAGACAGAAGAGCCCTTTTCCCTGGCTGTCAATCCCTGGAAGCAGGCACAGGGTGGGC 4680
Db ggaagacagaagagcccttcttccctggctgtcatlccctggaagcaggcacagggtgggc 4680
QY CATCATGAGACATANTGATCTCATCCCCCTTAAGAGCTGGCTGTACCTGAAGAGTATCAT 4740
Db catcatgagacataatgatctcatcccccttaagagctggctgtacctgaaaggatcat 4740
QY ACATGGCTGTGACCGCAAGATGGAGTGGCTGTCTATGGTGTGTGCCATCAACCCCTTCCA 4800
Db acatggctgtgacccgcaagatggaggtggctgtcatgggtgtgccatcaaccccttcca 4800
QY TTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCGCAGGCTGAGGGCTCAGCCAA 4860
Db ttgacctgcacacgcactegctgagcttctgcagctgcagcaggtgagggctgagccaa 4860
QY TGGGGCAGGACTGGCTAGGCCAGACTTGACTTGTCTGTGGGACCCCTGGGCAGGGGCACTTT 4920
Db tggggcaggactggctagccagacttgacttgcgtgtggagccctgggcaggggcactt 4920
QY TCCTTTCTCAGCTTCAGCTTCCCTCTCGGAAATAGGCTTAGTAATTCCTGGCCCTGGC 4980
Db tcccttctcagcttcagcttccccctctctggaaaaatgggtctagtaattcctggctggc 4980
QY CTTTCCAGGGCTCTTTGGGAGAGTAGAATTTGAGATGTGAATTTGCTTTGACTCATATAA 5040
Db ctttccaggctcttgggagatgaattgagatgtgaattgctttagactccattaaa 5040
QY GGGCTGTGCCAGAAATTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGTGGTCTGACCC 5100
Db gggctgtgccagaattttggcccttccacatggtgggtggctcctgttgggttctgacc 5100
QY CCACCTGTGCCCATAGGCTTAAGACCGCTTCCCTCCCTGTTCCGTTGGCTGCATACTCT 5160
Db ccacctgtcccatagctaggccaaaggaccgttctctctccctcttccgtgggtctaaactc 5160
QY CTCCTTGGGCTCTAGAGCTGTCTGGCTGTCTATGACTGGGACATCTGGAAGGTC 5220
Db ctctctggctctagaagctgtctggctgtctctatgacctgggacatctgaaaggctc 5220
QY AGTAGAGGAAGTGGCCAGGCTGGCCCTGTGTAGGCGGGGGGGCTGGGTGGCAGCCGTGAA 5280
Db agtagagggaagtggccaggctggcctggctggtagggccggggggctgggtggcagcctgaa 5280
QY TTATGATCTCTTCTAGGTACCCATGGCCCTTAGGGAACCTGGCAGATCTAGAGGAGCTG 5340
Db ttatgacctctctagttaccctcagcccttggcttgggaacctggcagatctaggaggagctg 5340
QY GAGCCCAACCCCTGGCGGCCGAGACCCACTCACCCCTCTTACCACAAGGTGGGGCATCTAAG 5400
Db gagcccaaccttgccggccagagcccaactcaacctctaccacaaggtgggggcatctaa 5400
QY GAGGGTCGAGAAGGAGACCCCTAACAGTGGCTGAGGCGAGGGGCCCTCATCTGGGCAGATG 5460
Db gaggggtgcagaaggagacccctaacagtggctgaggcaggggggccctcatctgggcagatg 5460
QY AGAAGAGAACTTTGTGTGTGGGGGTATCGCCCATCCAGCTCTCACCTTGTCTCAACTGT 5520
Db agaagagaaactttgtgtgttgggggggtatcgcccatccagctccacttctgtcaactgt 5520
QY GTCAGAAATCAGTTCAGTACGGGCTGTCTGAGGGGTGTCCAGGGTTCCCGAGGCTGGGAG 5580
Db gtcagaatacagttcagttcagtcaggtgtctgaggggtgtccaggggtccccagcctggag 5580
QY TGGCAGGGGTGCTATTTGTCCCTCAGCCCTGCTTTTGTGCCACGTCTTACTGTCTCTTC 5640
|||||

Db 5581 tggcagggggctgcatttgtccccctcagccctgcttcttttgcacatgcttactgtcttc 5640
QY CTGGAGTATTAACAGAGTCAAAATGTGTAGAGCACTGGAAGAGGGGTGTTCACTTTGGTG 5700
Db ctggagctaaacagaggtcaaaatgtgtgtaggagcacatgaagaggggtgtctacttgggtg 5700
QY GGTCTAGGTGGGAGGAGGCCCATTTGGGCTTGGCTTTGAAAGTCTTTGGTGTATGTAGAA 5760
Db ggtgtagggtggggagggagggccattgggtctgggtctgaaagatcttgggtgtagtagaa 5760
QY GAGTGTCTGAAAGAGAGAGGGGCCCTGAGCTCGAGGGGAGGCCCCACCCCTCGAGTCTG 5820
Db gagtgtctgaaagagagggggccctgagctcgagggcagggccccccctcagctgtg 5820
QY CCCAGGCCCTCAGCCAGCAGCTGTGTAGACCCAGGAGGAGAGCAGGTAGAACGGCTGGC 5880
Db cccaggcctcagccagcagctcctgttagaccagggagagacaggttagaaagggctggc 5880
QY AGCGAGTGGAGTGGAGTGGAGTGGAGAGTCCCTGGGATCTTCTGTGTGGCCCTTT 5940
Db agcgagtgagggtggggagtgagatggagagactcctcctggatcttctctgtggccctt 5940
QY CTGGGTGTGCCCTGGTGGGCATTTGTCCAGCAGGCGACGTGGGGCTGCTCCCTGAGG 6000
Db ctgggtgtgctctgggtggggcatttgtgccagcaggcagctggggctgctcctcctgag 6000
QY ATCTCTGTCCCTCACCTCCATCCAGGGCATTTGCCAGCAAGACCTACTATTCGGGATGAA 6060
Db atctctgctcactcactccatcccaaggcatgtgctcagcagaagacactactactcggagtagaa 6060
QY CACATCTACCCCTPACATGTATACCTGGCTGGCTTACCACTGTCTCGCAACCCCAATGTGCGGGAA 6120
Db cacatctaccctacatgtacctggctggctacacactgtcgcaaccgcaatgtgctgggaa 6120
QY GCCCTGCAGGCCCTGGCGGACACGGCCACTGTTCATCCAGAGTGCAGGATCCCCCTACTAG 6180
Db gccctgcaggcctgggcggacacggccactgtcatccaggagtggagatccccctactag 6180
QY GGCTGCAGCGCTGTCTCTTCTTCCCTCCCATCACTGTTTCCAAACACCCCTCTCTCCAGGACTG 6240
Db ggctgcagcctgtctcttcttccctccatcagtttccaaaccacccctgttccaggaactg 6240
QY AGGCCCTGGCTCCACGCCCCCATCCCCCTTCCATCCAGTCCCTAGGCAGCAAGGCCACCAT 6300
Db aggccctggctcccaagcggcccatcccccttccatccagtcctcaggcagcaaggccacat 6300
QY TACCAGGAGGTAGGACCCCTGATTAGGTGTACATCTTTTCCCTCTCCCTCTCCCTCTCCTC 6360
Db taccaggaggtagggaacctgattaaaggtgcacatcttctcctcctcctcctcctc 6360
QY CTAATTTTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATCTTAAACCACCATCATCCAGC 6420
Db ctatatttttttttctcagaaacagctcctaaatctccaaatgtttaaaccacacatcctcagc 6420
QY AGTGGGACTTCCACCCCTGGCCCCCATGCCCTCTCTCATTTCTTGTCTTCTTCTCTCTGGG 6480
Db agtgggacttccacctcggcccatgccccctcctcattcttcttcttctcctcctc 6480
QY CTGACCCAGACAGCATCATTTTGCAGTGCAGGACCCCACTACTTCCCCAGCCCCCTGGGG 6540
Db ctgaccagacagcatcttctgagtgaggaccccaactactccccccagccccctgggg 6540
QY CTCATCCCCCGCCAGGTCTCCCTGGGGCTACCCCGATGGGTGAGACCCCTTTCAGACCCCTAC 6600
Db ctccatcccccccgccagggtccccctggggctacccccgaatgggtgagaccccttccagaccctac 6600
QY AGAGACCCCACTGCTCTCACAGCTACAACCTACTTCCCGGGAAGCAGGAGATCTTACAAGG 6660
Db agagaccccaactgctctcagctacaaactactgcccggggaagcagagagatctcacaagg 6660
QY AGTTCTTTGAAGTAGCAATGATGTATCCCAACCTGTCTGAAGGAGGCACGACGCTTGC 6720
Db agttctttgaagtagccaatgatgtcatcccccaacctgctgaaggaggcagccagcttgc 6720

Qy 6721 TGGAGCGGGAGGAGCGGGGGAGCAAGCCAGGTGAAAGGCTGGAGCTCCAGCC 6780
Dbb 6721 TggaagcggcgagagcgccggggagcaaaagccaggTgaagagctgagactccaagc 6780
Qy 6781 TGTGTCCAGCTCCACCTGTGACAGAGGCTCCCTTCCACAGGGCCATGGGGGTGCATGTA 6840
Dbb 6781 Tgtgtccagctcccaacctccagacagggtcccttccacagggccatgggggtgcatgta 6840
Qy 6841 CGGGATTAGGATGCGACGAGGAGGTGGCCGTGAGCAGACAGCTATCTCCCTTTTCT 6900
Dbb 6841 cgggaattaggatggcaggaagaggtggccctgagcagacgatgttcccttttct 6900
Qy 6901 ATAAGTGAAGTCTGGGCGCCACGTTGGACGGGACTGAAAGGTATTTAGAGGTTTCTACCC 6960
Dbb 6901 ataactgaggtccctgggcccacgttgagcggagactgaaggtattttagaggtttctaccc 6960
Qy 6961 TGTGCTTTCAGTTTCATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGATGG 7020
Dbb 6961 Tgtgcttccagtctcatggccagactccctccctcagctgaggggtgaggtlaggagtg 7020
Qy 7021 TACGTCCTGGCTATGATGCTTTTATAAAGGAAGAGGTCTAAGAACTTCCCAACC 7080
Dbb 7021 Tacyccctggctatggatggctttataaaggaaagggttctaagaaatgttcccaacc 7080
Qy 7081 TATGCTTACCTTTTCTGAGGCAGGGGTCTTTTGCCTAGGTGGGGGCTGGCCTGTGCC 7140
Dbb 7081 tatgcttacctttctgagccaggggtctttgcttagtgaggggctggcctgtgcc 7140
Qy 7141 TCTGTAGGGGTGAGTGAAGAGACTGCTGTGCCCTCCCTTCCCTTCCCTGTCAGGGGAC 7200
Dbb 7141 Tctgtaaagggtgagtaagagactgctgtgcccctcccttccctcgtccagggcac 7200
Qy 7201 CCAGAGCAAGGTTCCGCGCTCCAGAGCCCTGAGTGTTCGCCACCTGCTGCGATTCTA 7260
Dbb 7201 ccagagcaagggttcggccctcccaaggaaacctgagTgcttcgcccacctgctgcgaattcta 7260
Qy 7261 CGACGGCATCTGCAAAATGGGAGGAGGCGAGTCCCACGCTGTGTGACAGTGGGCTGGGC 7320
Dbb 7261 cgaaggcatctgcaaaatgggagagggcagTcccaacgcgtgtgctgcacgtTgggctgggc 7320
Qy 7321 CACCTTTCTGTGAGTCCCTAGGCGCTTTTTCAGGGACAGGTGAGGGACAGCTGCACAGA 7380
Dbb 7321 cacccttctgtgcagtcctcctagggccgttttgagggacaggtgagggacagctgcacaga 7380
Qy 7381 GGTCTGGGCACTACAGTGGTGACAGCAGCCACGGCTTGTGACAGTCTTCTGGCGCCAGG 7440
Dbb 7381 ggtctgggcaactacaggtTgtgacagacagccagggcttgcagactttctggcccagg 7440
Qy 7441 GGCAGCATCTGCCATCCCTTTCGCTGCGATGGGACTGAGACCCCTTGGGTGGGATGGG 7500
Dbb 7441 ggcagcatctgccatcccttcggtgcgagtgggactgagacccctTgggtgggatggg 7500
Qy 7501 ATGGCCAGACGAGGCTCTGGAGTTCCAGCCACTGGCGCGCAACCTTCTCTACCTTGC 7560
Dbb 7501 atggccagagagggtctgtgagttcccgacctggccggcaacctgtctctcaactTgc 7560
Qy 7561 TCTCCCCACTGGCCAGGTGGGGCAAGAGTGGCATAGTCAGCCGAGAGGCGCAGGGCGG 7620
Dbb 7561 TctcccactggcccaggtTgcgcaagaagTgcgcaatagtgagccagagagccgagcg 7620
Qy 7621 CCGAGCGCAGAGCGGTGGGGCGAGGAAGCCCGGGAAGCGCGGCGGCGGCGCCACAGCC 7680
Dbb 7621 ccgagggcgagagcgTggggcgaggaagcccgggaaagccggcgcgggcccaagc 7680
Qy 7681 GGGAGTCCAGCCAGAGGCGCCCGCGCCCAAGAACGACGACTGGACAGGGGCTTGG 7740
Dbb 7681 gggagtccaagcagagagggcccgcccgcccaagaaagccagcactggacaagggccTgg 7740
Qy 7741 GCACCGCGCAGGGTGCAGTGTGAGGACCCCGCGGAAAGCCTCTGGGAGTGTGCTGGCA 7800
Dbb 7741 gcaaccggccagggTgcagtgtcaggacccccccggaaagccctcctgggactgtcgtggca 7800

Qy 7801 CAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCGACGATCACCAC 7860
Dbb 7801 cagcccgaggccctgaaggtggcagcagcgctcaggtgccagcaccgcagcatcaccac 7860
Qy 7861 CGCGGAGGCTCCAGTGTCTACTTTCCAGAGTGAAGATGAAGGCATGAAGGACTTGC 7920
Dbb 7861 cgcggagggtccagtgctcactttccagagtgaagaagtgaaggagcagtgaggagcgtgc 7920
Qy 7921 TGGTGGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGGCACAGTGCACAAG 7980
Dbb 7921 TggTggccaccacaagataactcagcgccaTcaagctgcaactcagcagcagtcgcaag 7980
Qy 7981 TGCAGATGAAGAAGCAGAAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCAAGC 8040
Dbb 7981 TgcagatgaagaagcagaaagtgtccacccctagTgactacactctgtcttcttccccaagc 8040
Qy 8041 GGCAGCGAAAGGCTCTGAAGTACTGGGGACTTCGGACCGCTTGTGGGGACCCAGGCTC 8100
Dbb 8041 ggcagcgaaaggcctctgaactactgggagcttcggaccgcttgTggggaacccaggtcc 8100
Qy 8101 CGCCTTACTGCCCAACTCTGAGCCCATGTTGTGCCCGCCAGCCCAAGGCGACAGGCTC 8160
Dbb 8101 cgccttagtcccccaactcctgagcccaTgttctgccccagccccaaaggagacagccctc 8160
Qy 8161 ACCTTACCCAAACCTAGGTTCCCGTCCGAGTACAGTCTGTATCAAAACCCACGATTT 8220
Dbb 8161 accttaccaaaacctaggTtcccggtTcccgagTcacagctgtatcaaaacccagattt 8220
Qy 8221 TCTCCAGCTCAGAACCCAGGGCTCTGCCCGCAGTGGTTAGAATATAGGTCTCTTCTCCAG 8280
Dbb 8221 Tctccagctcagaacccagggctctgccccagtgctgaaatataaggtctctcttcccag 8280
Qy 8281 AATCCACGCGGCCAATGAAACCTCAGGCTGGGTCTTAATTAACAGCTTTTAAAGGCC 8340
Dbb 8281 aatccacgcccgaatTgaaaacctcacgctgggtccctaaTtaccagctctttaaaggccc 8340
Qy 8341 AGCCCTTAGAACCCCAAGCTCTCTCGGAACCGCTACCTAGAGCCGACACCAAGCTTAC 8400
Dbb 8341 agcccttagaaacccaagctctctctcgaaacgctcctcctagagccagacaaagcttac 8400
Qy 8401 TCAGGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCAGGAGGACACAGG 8460
Dbb 8401 TcagggtctctcccagctgtTaggagctTgaggtTtcaaccttaacccaaaggagcgaagg 8460
Qy 8461 TCCCACCTCCAGCGCGGGAGCCTAGGACCACTACGCCCCCTTAGGAGTATATTTCCGCACT 8520
Dbb 8461 TcccaactccagcccgggagcctagTgagccactcagccctagTgagTataTtcttcgcact 8520
Qy 8521 TCAGAAATTCATATCTTGGGAATCCAAGCTCCCTGCCCCCAATAACTTCAGTCTCTTCT 8580
Dbb 8521 TcagaattccatatactTgcgaatccaagctcctcgcccccaataaactcagTcctgcttc 8580
Qy 8581 CAGAAATTTGGAATCCTAGTTTCTCTCTTCTGATCCGAGCTGCGGACACAAACATCC 8640
Dbb 8581 cagaaattTggaaatcctagTtctctctctctctctctctcagTatcccgagctTgggacaaaactcc 8640
Qy 8641 GCCCCAGCCTATGAGCATCTTCCAGCCCCGCGCTCTTCTTGACGAAACTTGGCCCCGGATC 8700
Dbb 8641 gcccccagcctatgacatcctTgagccccgcctctctctcagcaaaactTgcccccgatc 8700
Qy 8701 AGAGCAGACCTTCCCTTCCGACCCCTCTGGGAACCTTCCAGAGGTCCAGCCCATCTCCGAG 8760
Dbb 8701 agagcagaccttcccttccgacctctggaaacctcccagaggtccagTcrratctcggag 8760
Qy 8761 CATCCCGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAGCTGTATCTCTCTCT 8820
Dbb 8761 catcccgaggaataTctcgagaggggtTtagggagTgggTgacaagagccctTgaTctcttccct 8820
Qy 8821 GTTTTGTACATAGATTTATTTTTCAGTTCCAAAGAAAGATGAATACATTTTGTAAAAAAA 8880
Dbb 8821 gttttgtacatagatttttttctcagttcccaagaagatgaatacatattttgttaaaaaa 8880
Qy 8881 ATATAAAGCGCAAGTCCATGTTTATCTGGGAAATTTGGGGATGGGGGAGGAGTACGCGC 8940

Db 8881 atataaagcgaagtcacatgtttatctggaataattggaatgggagtggaagcgc 8940
QY 8941 CCCTTCTCCCTTTGCTTCTGGCTCCCGGACTTTTCGCTCCCTACTCTGTGGAGCGCA 9000
Db 8941 ccccttctccctttgtcttctgtcccggaacttttgcctccctacactgtggaagcgga 9000
QY 9001 GCGACAGTGGCGGGGAAGACGTAGGCTCCGCGCCCGCTTGGGGCTTCCCGCGCCCG 9060
Db 9001 gcgacagtggcgcggaagacgtaggctccgccccgaccttggggttcccccgcgcg 9060
QY 9061 CCGAGGCGCCGTCGCGCGGCGCTCTCCCGGACTGGCGGTGGGGCATCCNCGGCGCG 9120
Db 9061 ccgagggcccgctcccgggcgctctcccgactggcgactggcgatcccnngggcg 9120
QY 9121 GCCCGCGCCCGGCTTCAGCCCGCCCGCGGCTTCAGAGCCACGCGCGCCCGCCCG 9180
Db 9121 gccccgccccgggctcagccccgccccggcggttccagagccacggcgccccg 9180
RESULT 3
US-09-948-941-529
; Sequence 529, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; PRIOR FILING DATE: 2001-09-10
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 11190
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-529
Query Match 98.5%; Score 9040.4; DB 5; Length 11190;
Best Local Similarity 99.8%; Pred. No. 4.2e-169;
Matches 9165; Conservative 0; Mismatches 7; Indels 11; Gaps 11;
QY 1 CTGCTCTGAACCTCTGGCTCAGCAATCCTCTGCTTCAGCTTCCCAAGTGTCTTAA 60
Db 322 ctggctctgaactcctggcctcaagcaatcctcctgcttcagcttcccaagtggttaa 381
QY 61 TTACAGGCATGAGCGCTGGCATGAACCTTGACACTATTGAGATATACCTGGTCAAGTATTG 120
Db 382 ttacagcatgagcctggcatgaacttgacactattgagataatactggtcaggtatttg 441
QY 121 TGGAACTCCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGAGGAGATTATAA 180
Db 442 tggaaagtccctcaactctgtttggcagatgtttctcatgattagaggagattataa 501
QY 181 ATTTTGAAGAAATCCAGAGAGGTGAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
Db 502 attttgaggaaaatccagagaggtgaagaggtgaagtagggcagaaaatttaactgtttt 561
QY 241 ATTTTACTGCTATATPACCGAGTGTCTGGAACCTTGGCCCATGTTAGTACCAAAAATCTGTT 300
Db 562 atttaactgctataaccagagtgtctggaacttggcccatggttaagtaccacaaaatctgtt 621
QY 301 TTTTGTGAATGAATGAACAAATAAATGAGTGACCGGTGGAATTTAGTATTATTTTCAAGT 360
Db 622 tttttgaaatgaataagcaataaataatgagtgacctggaaatttagtattattttcaagt 681
QY 361 TTCAAAGCGTGTGATACAGGCGCAGGACAGTGGCTCACACCTGTAATCCACACATTT 420
Db 682 ttcaaagcgttgttgatacagggccagggcacagtggtgtcacacctgtaatccccacacttt 741

QY 421 TGGAGGCCGAGGTAGGAGGATCACTTGAGGTCAAGGAGTTCGAGACCAAGCTTGACCAACAT 480
Db 742 tggagggccgagglagagagatcaacttgaggtcagaggttcgagaccagcctgaccaaat 801
QY 481 GGTGACACCCCTGTCTCTACTAAGTAAATACAAAATTAGCCAAAGTGTGGTGCCAGGCA 540
Db 802 ggtgacacccctgtctctactaagtaaaatcacaaaattagccaaagtgtggtgagcgca 861
QY 541 CCTGTAATCCCGGCTACTTGGAACTGAGCGAGAAGTACACTTGAACCTTGGGAGGCGAG 600
Db 862 cctgtaatacccgctacttgggaagctgagcgagaagaatacaacttgaaccttggagagcag 921
QY 601 AGGTGTCAGTGAGCGGAGATCACCCCTGACCTCCAGCCTGAGTACAGAGCAGACTC 660
Db 922 aggttgcagtgagcgagatcacccctgcaactccagcctgagtgacagagcgagactc 981
QY 661 TGCTCTAAAACAATAAACAATAACTACTTCTTGGCCGGGTAAAGTGGTTCACGCGCTGT 720
Db 982 tgtctcaaaaacaataaacaataaactactcttttggcgggtaagggtgggttccagcctgt 1041
QY 721 AATTTTAGCACTTTGGGAGGCTGAGCGGCGAGATCACTTGAGGTTAGGGGTTTCGAGACC 780
Db 1042 aatttttagcaacttttggagggctgagggcgagatcaacttgaaggtttaggggttcgagacc 1101
QY 781 AGTGTGCCAACATGTTGAACCCCATCTCTACTTAAATACAAAAGTTTTCTGGGTGT 840
Db 1102 agtctgccaacatgttgaaacccatctctactataaatacaaaaagtttcttctgggtgt 1161
QY 841 GGTGGCGGAGCGCTATTAATCCAGCTACTTGGGACTTTTTTTTTTAAGCGGAATCTCACTC 900
Db 1162 ggtgcgagcgctataatcccaactcttgggactttttttaaagacgggaatctcactc 1221
QY 901 TGTGTCAGGCTGGAGTGGAGTGCGAGTGCGAAGATTCCTGCTCACTGAAGCTCCGCGCTCCAG 960
Db 1222 tgttggccccggctggagtgcaagttcttggctcactgagcctcccgccctccag 1281
QY 961 GTTCAAGGGGATTCGCCGCGCTCAGGCTCCCAAGTAGTGGGAATCCCTGCTCTCTGCAA 1020
Db 1282 gttcaa-gggattccccgcgctcagcctcccaagtagctgggaatccctgtctctgcaa 1340
QY 1021 AAAAAAATAAAAAAATAAAAAA-ATATATATATATATATATATATATATATATATATATAT 1079
Db 1341 aaaaaaataaaaaaataataataataataataataataataataataataataataataataata 1400
QY 1080 GTGTGTGTGTGTAT 1139
Db 1401 gt 1460
QY 1140 GGAGTGTGGCGCACGCTGTGATCACAGCTACTCGGGAGGCTGAGGCACGAGAAATCGCT 1199
Db 1461 ggaagtgtggtgagcagcctgtgatacagactactcggagagctgagcgagagaaatcgct 1520
QY 1200 TGAGCCCGTGAAGTGCAGGCTGCAGTGCAGCCCATGCAGGCGCTGCTGCTCCAGCTGGG 1259
Db 1521 tgagcccgtaagtgcaggtgcagtgagccagatgcagcactgcatccagcctggg 1580
QY 1260 CGAAAGAGAAAGACCGGTCTCAAAACAAACAAACAAAGCTACTTTAGCACCTGTGTAG 1319
Db 1581 cgaaaggaagagacggtgtctcaaaaacaaacaaacaaagctactctttagcacgtgttag 1640
QY 1320 AGTATCTCGCGGGGGAAGTGGGAACAGAGTGCCTGCACACAGAGTAGGCATCTTTTATATG 1379
Db 1641 agtatctcgggcggaagtgggaacagagtgctgacacagagtaggacatcttataatg 1700
QY 1380 TTACACACACTGATCCAGCTAAAGCGGCTGAACACATTTTACTCTCTGCGAGTGTGTTA 1439
Db 1701 ttaacagacactgataccacagctaaagcggtgaacacatttactctctgtgaggttata 1760
QY 1440 AAGTATCTGTTTTCATATTTTATTTTATTTTATTTTATTTTTCGATCAACCAACCTGAT 1499
Db 1761 aaagtaactgttttctcataatttttcttcttcttcttcttcttcttcttcttcttcttctt 1820
QY 1500 CTTTTTCTCTATAACTTGGCCGACCGCTGACAGCAAAACCGGCGAGAACTCGCGGAC 1559

Db 1821 ctcttctctcaaaacttcgacagaccgctgacagcaaaacgagacgctcgac 1880
Qy 1560 CTCACACCCGAGTCTGCAGGTAGTCCCGGACTACATTTCCAGAGCACTGCCGG 1619
Db 1881 ctcccaccacgctgctgaggtgagtgccgagcagctatcttccagagcactgcg 1940
Qy 1620 CACGCTTCCTGCTGCTGGCTGAAGGAAAGGGCAATCCCTGAGTATCTCGGGAAGA 1679
Db 1941 caagcttccgctgctgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 2000
Qy 1680 GGTGTCCGAGCGCGGACCTAGAGATCCAGAACCCACAGCGGCGCGGCGCCGCC 1739
Db 2001 ggtgtccgagcgcgcgacccatagatccagagccagcgcgcgcgcgcgcgcgc 2060
Qy 1740 ACTATTTCAGGCTCTGCGGCGGAGGGGTGGGCCGAGACTCCACTCCCGGGGTAGTG 1799
Db 2061 actattccagagctctgcgagcaggtgaggtgaggtgaggtgaggtgaggtgag 2120
Qy 1800 CGACCTTAGGGCGGAGCTTTCATGTCCAGCAGGCTCCGGCGGCTCGCGCGGCTGCC 1859
Db 2121 gacccctagggcgagcttcatgttccagcaggtctccgagcgcgcgcgcgcgcgc 2180
Qy 1860 TAGTGTGGATGTAAAGCGGAGGTGGCGGAGGGGAGCCGAGGCACTCTCTTGGG 1919
Db 2181 tagtgtggaagtgaagcgcgaggtgaggtgaggtgaggtgaggtgaggtgaggt 2240
Qy 1920 GTTTCGGGCTGACCTGGGTGCGCTTCTGACAGACTTTACAGCCCGGGGACACAG 1979
Db 2241 gtttgggggcttgacctgggtgaggttctgagagacatttacag-ccccgggggacag 2299
Qy 1980 TCGTAGAGAGGGCGGGCGGCCATTGGGGCTCTCTATGGGTGCTTGGGGCGCACCC 2039
Db 2300 tcgtagagagggcgcgcgccattggtggtctctcatgtgggtggtggtggtggtggt 2359
Qy 2040 CATCGGCTACCGGGCTCCCGGAATTGTGGGGACAAAGAGCTCTGCAGTCTCGCTGA 2099
Db 2360 catcggttaccggcgctcccggaattgtgggggacaaagagctgagctcgcgctgga 2419
Qy 2100 GGGCTCTACCCACAAAAGAGGGGAAGCGGTGAGCAGAGCTGAAGAGGTGGGGAAGC 2159
Db 2420 ggggtctcacgcacaaagagaggggaagcgggtgagcagaggtgagaggggtgag 2479
Qy 2160 AGGGAGTGTGCTGTGTGCGGGCGGGTGGAACTTACGGAGCCCTGGGAGAGGCTCC 2219
Db 2480 agggagctgtgctgtgctgagggcggtggaaccttagcgagacctgggagaggtccc 2539
Qy 2220 CCGGCGGAACCTGCGGACCTCCCTCCCGGGCTTGCCTTGCAGGCGCGCGCCACCGC 2279
Db 2540 ccggcgaaacctgcccgaacctccctcccccggcttgccttgcagggcgccgcccacgc 2599
Qy 2280 CCGCGCGATGGGGCTGAAGGCGCGCAGAGACGCTGTTCGCGTGCCTCCATCGAGC 2339
Db 2600 ccgcgccaatggggcggaagggcgcccgagagagcgtgttcccgctgcgctccacg 2659
Qy 2340 AGTGGTGGCTGTTGCTGCTGCGAGCTGGGCGGAGAGCGGACCTGTGCTCTCTTT 2399
Db 2660 acgtggtgctgctgttctgctgcgagctggggcgagagagcgagacctggtgctccctt 2719
Qy 2400 CCTTGGTGTGGCTTGTGTGAGCATTTTCTGGCTGTCAACGGCGCTCATCCCTACCAAG 2459
Db 2720 ccttgggtggtgctgtgagcatttctggctgtcaacggcgctcatccataccacg 2779
Qy 2460 TTCCCGAGCTACCTTCAGCCCGCGCGCCCGAGCGGCTGGCGGCTTCACCTACT 2519
Db 2780 tccccagctcaaccttcagagcccgcccccgcccccgcccccgcccccgcccccg 2839
Qy 2520 TTCCCGTGGCGGACCTGTATATCATCGCGCGCTCTATGCCGCTTACCGCGGAGATCC 2579
Db 2840 tccccgtggcgacctgtctatcatcgcgcgcctctatgcccgttcaacggcgctccacg 2899
Qy 2580 GAGGCGCGCTGACCTGTCCCTCTATCTCGAGAGGGGGTGTCTCCAGCGGTGAGCTGG 2639
|||||

Db 2900 gaggcgctgacactgtccctctatactcagagaggggtgtctccacccgtgagctgg 2959
Qy 2640 TGAAGAAGGTCTCCGATGTATGGAACAGCCTCAGCGCTCTACTTCAAGTATCCGG 2699
Db 2960 tgaagaaggctcccgatgtcatatggaacagcctcagcgcctccctactccaaggatcg 3019
Qy 2700 CCCACATCCAGTCCCTCTTACAGTTCATCAGAGTTGGAGCCCACTAGTGTGGAACTTA 2759
Db 3020 cccacatccagctccctctcagcttcatccacaggttgagggccagtagtgaggatctta 3079
Qy 2760 TCCATGACCCACTTCTTCAAAACCCCTCCATGGTTTACAGAACCCCTTTTAAAGACTGTAA 2819
Db 3080 tccatgacccacttctcaaacccctccatggtttacagaaaccccttttaagaactgtaa 3139
Qy 2820 CTTGTGAGGTTCGGCAGGTGTATTTTCTTTCAGTTGGGAAACTGAAGCCACAG 2879
Db 3140 cctgtgaggttcggcaggtgttatcttctcttctgagtttgggaaactgaaagccagag 3199
Qy 2880 AGGGAAATGATATGCCAAAGTTCACACAGGCTGGCAGGCTGGAAGTGAAGCTGTATC 2939
Db 3200 aggggaatgatatgccaaagtccacacgcgcaggtgagggcgtgaggtgagggcgtgac 3259
Qy 2940 ACTTGGCTCCAAATCATCACTCAGCTTGCCTCCTCAGCACCCACCTTGGCAGCTG 2999
Db 3260 acttggctccaaatcatcaacctcctctgccccctcagcaccccccccttgcacctg 3319
Qy 3000 AACAGCTACAGAGTTCCTAAGCATGAGACAGAGAGGGCGGACAGAGATTTAGGGGGCAAG 3059
Db 3320 aacagctacagagttctlaagcatgagacacagagggcgagcagagattttagggggcag 3379
Qy 3060 AAGTGAATTTGGGCTGCTATTTGAGGCAAGTTTAAACAAATATATGCTATGAAGATTTT 3119
Db 3380 aagaagaaattgggtgctgcatgtgagcgagtttaacaaataatggctatgaaagattctt 3439
Qy 3120 TTTTGT 3179
Db 3440 ttt 3499
Qy 3180 GATCATGTGCTCACTGCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACCTCAGCCTC 3239
Db 3500 gatcatgctcactgcagcctcagctccctgggtcagagatctcccaacctcagcctc 3559
Qy 3240 CTGAGTACCTGAGATCAGGATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3299
Db 3560 ctgagtagctgagatgacagcagctgagcagctgagctgagctgagctgagctgagct 3619
Qy 3300 GTAGAGATGCTGCTCACTATGTGGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3359
Db 3620 gttagagatgggtgctcactatgtggccagagctgcttgaactcttgggtccagctgac 3679
Qy 3360 CTGCCCGCTCAGTCTCCCAATGCTGGGATTTACAGGTGTGAGCCACCGCAACTGCTGGC 3419
Db 3680 ctgcccgcctcagctcccaaatgctggatgaggtgagccacccgcaactggtgagc 3739
Qy 3420 CTATGAAATTTTGT 3479
Db 3740 ctatgaaattt 3799
Qy 3480 AGTGGTGAATCTCGGCTCACTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3539
Db 3800 agtgggtgaaatcctgctcactgcaagcttctgctccttctcctgctcctcctgctc 3859
Qy 3540 TCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGAGGCTGCCACCACTGCTGCTGCTGCT 3599
Db 3860 tctgctcagcctcctgagtagctggagctacagagagctgcccacclgctggctaat 3919
Qy 3600 TTTTGT 3659
Db 3920 ttt 3979
Qy 3660 TCCTCAGCTGCTGATCCCGCTTGGCTCCCAAGTGTGGGATTTACAGGCTGAGC 3719
Db 3980 tctgacacctgagatccgcccccttggctcccccaagctgctgggttaccagggctgagc 4039
|||||

Qy	3720	CACGCGACCTTGGTCAAAAATGTTTGAGACAGAGAAGGGGCTTGACCTCAAAAAGCTTTAAG	3779
Db	4040	cacgcaccttggctcaaaaaatgtttgagacagagagggccttgacctcaaaaggcttaag	4099
Qy	3780	AGTCAGGGCTTGCAAAAGACTTTGACCAAGCCCGTTGACTGCGCAATCCATCCTGGTG	3839
Db	4100	agtcaggccttgcaaaagactttgacaaagcccggcttgactggcaatccccactcctggt	4159
Qy	3840	TGCCATATTGAGAAGCAATCAGAGGCTGCTCTCAGCTTAGCAGSAAAAGAGTCACAGAGA	3899
Db	4160	tgccatatgagaaggaatcaagaggctgctctcaagcttagcagagaaaagaagtgcagaga	4219
Qy	3900	TAAATGAGGGTTATTTGTTGGTGGGTGATAGCCAGAGAGTGTGTGGCCAGCGTCCCTGTTT	3959
Db	4220	taaatgagggttatcttctgttgg-gggtgtatagccagagagtgttggccagcgctcctgttt	4279
Qy	3960	TTGCCATCTCTGTTTTTAACCTAGTAGTCAGTAAATGGAATCCCTTAAATCCATAGAAAT	4019
Db	4280	ttgccatctcctgtttcaacctagtagtgcagtaaaatggaaatccctaaatccctagaaat	4339
Qy	4020	ATATAATFAGAGTTTGACAGAAAGACAGAGTAGGTAGGCGCAAGGCTGGGTACAGTACAGGATA	4079
Db	4340	ataataatagagtttgcagagaaagacagagtagggccaaaggcttggctcagctacagagata	4399
Qy	4080	TCCAGAAAGTATCTTGTGGACATAGAGGGTGTAAACAGGGGAGAGAGTCTTTGAACACGC	4139
Db	4400	tcagaaaagtatcttctgttggacatatagagggtgttaaacagaggagagagcttttgaacacg	4459
Qy	4140	TGGAGGGAAGGATGAGGGATAGTGGCGCAGAGAAATCTGAGTGTGGGTACACAGCTTG	4199
Db	4460	tggagggaaaggatgagggatagtgggcagaggaatctgaggttgggtcacagggcttg	4519
Qy	4200	GAAAGGAGTGGCAGGAGTGTGGCCATCACTACTCTGG-CCCCCTTCCCATGCTTAAAG	4258
Db	4520	gaaaggagtgggagggagtg-ggcccaatcaatactctggcccccttccccatgtttaaag	4579
Qy	4259	CACAGAGACCCCTTTCATTACCTCCCCCTTCCACAGGCACCAAAATGGACAGCTCCGG	4318
Db	4580	cacagaggacctcttcttcaataacctcccccttccacagggaccaaaattgacacagctccgg	4639
Qy	4319	TGTGGCCTTTGCTGTGGTGTGGGGCTTGCAGGCCCTGGGTCTCGGGATGTCACCTCCG	4378
Db	4640	tgtggccttctgtctggttggggcctgcagggcctgggtctccgggagtg-ccccactctgc	4699
Qy	4379	CCTGCTCTGAGATCATGCTTGGTGTGGTGGCCCAATGGGAGCAGACAGCTGAGGT	4438
Db	4700	cctgtctcgaggaatca-ctgcttgggtagtg-cttggggcccaatggggagcagctgaggt	4759
Qy	4439	CACCTGCGACGCAAGGGCAACGAGACCGCAGGGGCCAGACAGTCAATGCGCGTGTGGC	4498
Db	4760	cacctggcacgcaagggcaagaggaacgagggccagacagctcaatgcccgtgtggc	4819
Qy	4499	TGACGGGTATGTTCCCTCCGCCAGCCTTGTCCCTTCATCTGCTAGTGTAGTACGCCAAGCC	4558
Db	4820	tgagcgggtattgttccctccccacagcctgttcccccttcaatactgtagtagccccagcc	4879
Qy	4559	ACCCAAGGACTCCATTCTTGGGCCACACCCCTTCTTCCCATCACCCACCCACATAGG	4618
Db	4880	acccaaggactccattcttcttggggccaccccccttcttcccaatcaacccacacataagg	4939
Qy	4619	AAGGGAAGACAGAGAGCCCTTTTCTGGCTGTCTATTCCCTTGAGCAGGACAGCGGTGG	4678
Db	4940	aagggaagacagaagagcccccttttcttggctgtcatctccctgaagcagcaagagtggt	4999
Qy	4679	GCCATATGAGACATATGATCATCCCCCCCTTAAAGAGCTGGCTGTACCTGTGAAGGATC	4738
Db	5000	gccatcatgagacataatgatctcaatccccctcaagagcttggctgtaccttgaaggatc	5059
Qy	4739	ATACATCGCTGTACCCGCAAGATGAGGTGGCGTTCATGGTGTGTGCTGCCATCAACCCCTC	4798
Db	5060	atacatcgctgtgacgcgaagatggaggtggcgcttcatgtgtgtgtgccaatcaacaccttc	5119

Qy	4799	CAATTGACCTGCACACCGACTCGCTGAGAGCTTCTGACGCTGACGAGGTGACGAGGTGAGGCCTGAGCC	4859
Db	5120	cattgacctgcacacgactcgtgagctctcagctcagcaaggtcgagggctgagcc	5179
Qy	4859	AATGGGCGAGGACTGGCTTAGGCCAGACTTGACTTGCCTGGGACCCCTGGCAGGGGCAC	4918
Db	5180	aatggggcaggaactgggctagggcagaacttgacttgccttgtagggacccctggggcagggcac	5239
Qy	4919	TTTCCCTTCCTGAGCTTCAGCTTCCCTCCTCGAAAAATGGGTAGTAAATTCCTGGGCTG	4978
Db	5240	tttcccttctgagcttcagcttccctcctcgaaaaatgggttagtaattcctgggctg	5299
Qy	4979	GCCTTTCCCGAGGCTCTGTGGAGAGTAGAATTGAGATGTGAAATTCCTTGCACCTCAATTA	5038
Db	5300	gcctttcccgaggctcttgggagagtagaatttgagatgtgaaattgctttgactccatta	5359
Qy	5039	AAGGCTTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTTGGTCTCTGAC	5098
Db	5360	aagggtctgttcccaagaatttggcccttccacatggtgggtgggtccctgttggttctcgac	5419
Qy	5099	CCCCACCTCTGCCCGTAGGCTAAGGACCCGTTCTCTCCCTGTTCCCTGGCTGACTATAACT	5158
Db	5420	ccccactctgcgcgataggttaaggaccggttctctcccttgttccgtgactcataaact	5479
Qy	5159	CTCTCTTCGGCTCCTAGAACTGCTCTGGCTGCTCTATGACCTGGGACATCTGGAAGG	5218
Db	5480	ctctcttggctcctagaagctgctggctgctctatgacctgggacatctgaaagg	5539
Qy	5219	TCAGTAGAGGGAAGTGGCCAGGCTGGCCTGGTGGAGCGCGGGGGCTGGGTGGCAGCCGTG	5278
Db	5540	tcagttagggaaagtggcaggctgcgcttggtagggccgggggctgggttggcagctg	5599
Qy	5279	AATTATCATCTTTCTTAGTATCCCATGGCTTAGGAACTGCGACATCTAGAGGAGC	5338
Db	5600	aattatgatcttctctagtagtcccatggctcttgggaaccttgcagatctcagaggagc	5659
Qy	5339	TGGAGCCACCCCTGGCGGCGCCAGACCCACTCACCCCTCTACACAAAGTGGGGGCATCTA	5398
Db	5660	tggagccacccctggcggccagaccactcaectcaecttaccacaaggttgggggcatcta	5719
Qy	5399	AGGAGGTGCAGAAAGGAGACCTTACAGTGGCTGAGCGCAGGGGGCCCTCATCTGGGCAGA	5458
Db	5720	aggagggtgcagaagggagaccctaacagtggctgaggcaggggccctcatctggggcaga	5779
Qy	5459	TGACAAGAGAACTTTGTGTTGGGGGTATCGCCATCCAGCTCTCACTTTGTGTCAACT	5518
Db	5780	tgaagaagaactttgtgtgttgggggtatcgcccatccagcttcaecttgtgtlcaact	5839
Qy	5519	GTGTGCAAGATCAGTTGCTCAGTCAGGGCTGTCTGAGGGGTGTCCAGGGTTCCCCAGCCCTGGG	5578
Db	5840	gtgtgcagaatcagtlcagtlcagggtgtctgaggggtgtccagggttccccagccctggg	5899
Qy	5579	AGTGGCAGGGGCTGCATTTGTCCTCCCTCAGCCCTGCCTTTTCTGCCACTGCTTACTGTCTCT	5638
Db	5900	agtggcagggttgcatttgcctccagccctgctcttcttgcacttgcctactgtct	5959
Qy	5639	TCCTTGGAGTATACAGAGGTCAAACTGTTAGGACACTGAAGAGGGGTGTTCACTTGG	5698
Db	5960	tccctggagtagtaacacagaggccaattggtaggagcacgaaagagggggtgttcaacttgg	6019
Qy	5699	TGGGTGTAGGTGGGGAGGAGGGCCATTGGGCTGGCTTGAAAGTCTTTGGTGA TGCTAG	5758
Db	6020	tgggtgtlaggltggggaggaggccatttggccttggaggttcttggagagtttltggltgtgtag	6079
Qy	5759	AAGAGTGTCTGAGAAAGAGAGAGGGCCCTGAGCTCGGAGGGCAGGCCCCACCCCTGCACTG	5818
Db	6080	aagagtgtctgaaaaagaaaggccctgagctcgaggcaggcccccaccctcgactc	6139
Qy	5819	TGCCCCAGGCTCAGCCAGAGTTCCTGTAGACCCAGGGAGGAGACAGTAGAGGGGCTG	5878
Db	6140	tgcgccaggctcagccagcagtcctctgtagaccagggaggagacaggtagaagggtg	6199
Qy	5879	GCAGCGAGTGGAGTGGGAGTGGAGATGAGAGAGACTCCCTGGGATCTTCTGTGGGCC	5938

||||| 6200 gcagcgagtgaggctggaggtggagatgagaggactccctgggatcttctgtggcccc 6259
QY TTCTGGGTGTCCTCGTGGGGCATTTCTGCCACGAGCGACCTGGGCTCCCTCCCTGA 5998
Db ttctggggtgtgccccctggctggggcatgtgtgcagcagggcagctggggctgccccctga 6319
QY GGATCCTGTGCTCACCTCCATCCATCCAGGCGATTCCTCAGCCAAAGACCTACTATCGGGATG 6058
Db ggatcctgtcctcaactccatccatccaggcatgtgctcagccaaagcctactatcgggatg 6379
QY AACACATCTACCCCTPACATGTACCTGGCTGGCTACCACTGTGCAACCGCAATGTGCGGG 6118
Db aacacatcaccctacatgtacctggctggctaccactgtcgaaccgcaatgtggcgg 6439
QY AAGCCCTCGAGCCTGGCGGACAGCGGCACTGTCCATCCAGGAGTGGAGTCCCCCTACT 6178
Db aagccctcgagcctggcgagacggcgaactgtcatccaggagtgaagatccccctact 6499
QY AGGGCTCGAGCCTGTCTTTCTTCCCTCCATCAGTTTCCAAACACCCCTCGTCAGGAC 6238
Db agggcctcgagcctgtcttcttccctccatcagtttccaaaccacccctcgtccaggac 6559
QY TGAGCGCTGGCTCCACGCCCCATCCCTTTCCATCCAGTCCCTAGGCGAGCAAGGCCACC 6298
Db tgaggcctggctcccaagccccctcccttccatccagtcctcagtcagcagaagccacc 6619
QY ATTACCCAGGAGTAGGACCCCTGATTAAAGGTGTACATCTTCCCTCCCTCCCTCC 6358
Db attacccaggagtaggacccctgattaagggtgtcaacttcttccctccctccctctcc 6679
QY TCCTAAATTTTTTTTCTCAGAACAGTCTCAAACTCCCAATGTTTAAACACCATCATCCA 6418
Db tcttaatttttttttctcagaacagtcataaattccaaatgtttaaccacacatccca 6739
QY GCAGTGGAGTTCACCTCGCGGCCCATGCCCCCTCTCATCTTGTCTTTCCTCTG 6478
Db gcaagtggagacttccacccctggccccctgccccctctctctctctctctctctctg 6799
QY GCCTACCCAGACAGCATANTTTGACGTAGGACCCCGACCTACTCCCGCCCGCTCGG 6538
Db gctcaccagacagca tcaatttgcagtgaagacccacacactactcccccgccccctggg 6859
QY GGCTCCATCCCGCCAGGTCCCTGGGCTACCCCGCATGTGTAGACCCCTTCAGACCT 6598
Db ggctccatcccccgccaggtccctgggctaccctcccgatgtgagacccctcagacct 6919
QY ACAGAGCCCTACTCTCTACAGCTACAATCTGCGCGGGAAGACGAGGAGATCTACAA 6658
Db acagagacccactgtctcagcagctacaaactactgcccgggaagacgaggagatcaca 6979
QY GGAGTTCTTTGAAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGAGCCAGCTT 6718
Db ggaagtctcttgaagtagccaa tga tgcatacccccaacctgctgaaggaggcagcagctt 7039
QY GCTGAGGCGGCGAGGAGCGCGGAGCAAGCCAGGTGAAGGCTGGAGTCCAG 6778
Db gctgagcgggcgagagcgcgcggggagcaagccaggtgaaaggctggagctccag 7099
QY CTTGTGTCAGCCTCCACCTGAGAGGGCTCCCTTCCACAGGGCCATGGGGCTGTCAG 6838
Db cctgtgtcagcctcccaactggacagggctcccttccaca -ggccatgggggctgcag 7158
QY TACGGGATTAGGGATGGGAGGAGGAGTGGCCCTGAGCAGACAGTATGTTCCCTTTTG 6898
Db tacgggattcagggatggcagggaggaggtggccccctgagcagacagctatgttccccctttg 7218
QY CTATAACTGAGTCTGGGCCCACTGTGACGGGACTGAAGGTATTTAGAGGTTTCTAC 6958
Db ctataactgagtcctggccccacgttggagcgggactgaaggtatttttagaggtttctac 7278
QY CTTGTGCTTCAAGTTTCATGGCCAGACTCCCTCCCTCAGCTGAGGGGTAGGGAT 7018
|||||

Db cctgtgcttccagtttcatggccagactccctccctcagctgaggggtggaggtcagggt 7338
QY GGTACGTCTGCTGCTATGATTGGCTTTATAAAGAAAGAGAGGTCTTAAGAAATGTTCCCAA 7078
Db ggtacgtctcgtcgtatggatttggctttataaaagaaagaggttctaaagaaatgtlcccaa 7398
QY CCTATGCTTACTTTTCTGGAGCCAGGGGTCTTTGCCCTAGGTGGGGGCTTGGCTGTGTC 7138
Db cctatgcttactcttcttctggagccaggggtcttgcctcaggtgaggggctcgtgctg 7458
QY CCTCTGTAAGGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTCCCTCGTCAGGGC 7198
Db cctctgctaagggttgagtaagagactgactgtgcctcccttccctccctcgtccagggc 7518
QY ACCCAGAGCCAAAGTTCGCCCTCCAGGACCTGAGTCTTGCCTCCCTGCTGCGATTTC 7258
Db acccagagccaaaggttcccgccctccagaccctgagtgcttgcgccacctcgtgcgactc 7578
QY TACGACGSCATCTGCAATGGGAGGAGGCGAGTCCCAGCCTGTGCTGCACGTGGGGCTGG 7318
Db tacgacggcatctgcaaatgggagaggcagtcaccaagcctgtgctgcagctgggctgg 7638
QY GCGACCTTTCTTGTGCAGTCCCTAGGCGCGTTTGGAGGACAGGTGAGGAGACAGCTGCACA 7378
Db gccaccttcttctgtagtccctcagggcgttttgaggacaggtgagggacagctgcaca 7698
QY GAGGTCTGGGCACTACAGTGTGTGACAGCAGCGGCTTGTGAGACTTTTCTGGCCCA 7438
Db gaggtctggggcactacaggttggtagacagcagccagggctgtcagagacttcttgcccc 7758
QY GGGCAGCATCTGCCCATCCCTTCTGGTGGCGATGGGACTGAGACCCCTGGTGGGATG 7498
Db gggcgagcatctgccccctccctcgggtgcagatgggagctgagacccctggggtggatg 7818
QY GGATGGCCAGAGCAGGCTCTGGAGTTCACGACCTGCGCGGCAACCTTGTCTCACCTT 7558
Db ggaaggcagagcaggggtccctggagttccagccactgcccggcaacctgtctcacctt 7878
QY GCTCTCCCTACTGGGCCAGGTGCGGCGAGAGTGGGCTAGTGAGCCGAGAGCGCCAGGC 7618
Db gctctccccactggccccaggtgcggcagaaaggtgcgcatagtgagccagagagcgagcg 7938
QY GCGCAGGCCAGGAGCGCTGGGCGAGGAAGCCCGGAAAGCCGGCGCGGCGGCGCCACG 7678
Db ggcgagggcgagcgagcgctggggtgcgaggaagcccggaagcgcggtggcgccacg 7998
QY CGGGAGTCCAGCCAGAGGAGCGCCCGCCCGCAAGAGCCAGCAGCTGGACAGGGCT 7738
Db gggggagtcctcagcgagggagggcccccgccccaggaagccagcactggacaagggcct 8058
QY GGGCAGCCGCGCAGGCTGAGTGTACGAGACCCCGGGAAGCTCTCTGGGACTGTCTCGCTGG 7798
Db gggcagcgccaggggtcagtgctcagagcccccggaagcctcctcgtggactgtcgctgg 8118
QY CACAGCCGAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 7858
Db cacagcccgagggcctgaaggtggcagcagcgtcaggtggcagcagcccgagcagctacc 8178
QY ACCCGCGAGGCTCAGTGTCTACTTTCCAGAGTGCAGAGATGAAGGCGATGAAGAGCT 7918
Db acccgaggggtcagtgctcacttccagagtgagaga tgaaggagcag tgaaggagccl 8238
QY GCTGTGGCCCAACAGATCACTCGAGGCGCATCAAGCTGCAACTGCAAGCTGAGCTGCACA 7978
Db gctggtggccaccaaga tcaactcgagccatcaagctgcaactcagcagcagctgcag 8298
QY AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTCTTCA 8038
Db agtgagatgaagaagcagaagtgctccacccctagtgactacactctgtcttctctcaa 8358
QY GCGGAGCGCAAGGCCCTCTGAACCTACTGGGACTTGGAGCGCTTGTGGGACCCAGCG 8098
Db gggcgagcgcaaaagccctctgaactcctggggacttccggaccgctgtggggaccacagcg 8418

Qy	9178	CCG 9180	
Db	9494	ccg 9496	
RESULT 4			
US-09-948-941-649			
; Sequence 649, Application US/09948941			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL000788			
; CURRENT APPLICATION NUMBER: US/09/948,941			
; CURRENT FILING DATE: 2001-09-10			
; PRIOR APPLICATION NUMBER: 60/231,328			
; PRIOR FILING DATE:			
; NUMBER OF SEQ ID NOS: 13618			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 649			
; LENGTH: 11205			
; TYPE: DNA			
; ORGANISM: Human			
US-09-948-941-649			
Query Match 98.5%; Score 9040.4; DB 5; Length 11205;			
Best Local Similarity 99.8%; pred. No. 4.2e-169;			
Matches 9165; Conservative 0; Mismatches 7; Indels 11; Gaps			
Qy	1	CTGGTCTTGAACCTCCTGGCCCTCAAGCAATCCTCTGCTTCAGCTTCCCAAGTGTGTGTA	60
Db	322	ctggctctgaactcctggcctcaagcaatcctcctgcttcagcttcccaagtgctgtaa	381
Qy	61	TTACAGGCATCAGCCTGCATGAACCTTGACACATTTTCAGATATACTCGTCAGGATTTTG	120
Qy	382	ttacaggcatgagcctggcagaacttgacactattgagatactactggtcaggtattcttg	441
Qy	121	TGGAATGTCCTCAACTCTGTTTTGCCAGATGTTTTCTCATGATTAGAGGAGAGTTATAA	180
Db	442	tggaaatgccctcaactctgttttgccagatgtttctcatgattagagaggagtcataa	501
Qy	181	ATTTTGAGGAAATCCAGAGAGGTGAGAGGTGAAGTGAAGTCAGGCACAGAAATTTAATCTGTTTT	240
Db	502	attttgaggaataatccagagaggtgaagtgaaagtagggcagagaataatlaactcgtttt	561
Qy	241	ATTTACTCTATATACCGAGGTCTGGAACCTTGCAACCTTGCCCATGCTAGTAGTACCAAAATCTGTT	300
Db	562	atttactctctataaccagagtgctgtgaacttggtcccatggtaagtcaccagaatactcgtt	621
Qy	301	TTTTTTTGAATGAATAAGCAAAATAATGAGTCGACCGTGGAAATTTAGTATTATTATTTCAAAGT	360
Db	622	ttttttgaaatgaataagcaataaataatgagtgaccgttggaatttagtattatttccaagt	681
Qy	361	TTCAAGCGTTGTTGATATACAGCGCAGGCACAGTGGCTTCACACCTGTAAATCCCAGCACTTT	420
Db	682	ttcaagcgttgttgatacaggccaggccacagtggtccacactgtaatcccagcacttt	741
Qy	421	TGGAGCGCGAGGTAGGAGGATCATTGAGGTCAGGAGTTTCGAGACACAGCCTGCACCAACAT	480
Db	742	tggaggccgaggttaggaggtaccttgaggtcaggagttctgagaccagcctgaccacaat	801
Qy	481	GGTGACACCCCTGTCCTTACCTACCTAAATTAACAAAATTAGCCAAAGTGTGCTGGCAGGCA	540
Db	802	ggtagacacctctgtctctactaagtaaaaatacagaataatagccaagtggtggcaggcca	861
Qy	541	CCTGTAATCCCGCGTACTTGGGAAGCTGAGGCAGGAAGTAATCACTTGAACCTGGAGGACG	600
Db	862	cctgtaaatcccggtacttgggaagctgaggcagaagaataacttgaacctggaggagcag	921
Qy	601	AGGTTGAGTGAAGCGGAGATCACCCCACTGCATCCAGCCTGAGTGACAGAGGAGACTC	660
Db	922	aggttgagttagccgagatcaccccaactgacttccagccttgactgacagagcagagctc	981

QY 661 TGTCTCAAAACAAATAAATAACTACTCTTTGGCGGGTAAGGTGTTTACACGCTGT 720
|||||
Db 982 Tgtctcaaaacaaaataaactactcttggcgggtaagtggttcaacgctgt 1041
QY 721 AATTTTAGACATTTGGAGGCTAGCGGGCAGATCACTTAGAGTTAGGGTTCCAGACC 780
|||||
Db 1042 aatttttagcaacttggagctgagcgcgagatcaacttgaggttaggggttagagacc 1101
QY 781 AGTCGTGGCCACATGTTGAAACCCCATCTACTTAAATACAAAAAGTTTCTGGGTGT 840
|||||
Db 1102 agtcggcgaacatggtgaaaccccatctactcaatacaaaaagtttcttgggtgt 1161
QY 841 GGTGGCGGACGCTATAATCCAGCTACTTGGGACTTTTAAAGAGGAATCTCACTC 900
|||||
Db 1162 ggtggcgagcctataatccagctacttgggaacttttttaagagaggaatctcactc 1221
QY 901 TGTTCGCCAGGCTGAGTGCAGTGGCAAGATCTGGCTCACTGAAGCCTCCGCCCTCCAG 960
|||||
Db 1222 tgttggcaggctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1281
QY 961 GTTCAAGGGATTCCCGCCTCAGCTCCTCCAGTACCTGGGAATCCCTGCTCTGGAA 1020
|||||
Db 1282 gttcaa-999atccccgcgcctcagctcccaagctagctgggaatccctgtctctgcaa 1340
QY 1021 AAAAAAATAAAAAAACAACAAA-ATATATATATATATATATATATATATATATAT 1079
|||||
Db 1341 aaaaaaataaaaaaataataataataataataataataataataataataataataata 1400
QY 1080 GTGTGTGTGTAT 1139
|||||
Db 1401 gtgtgtgtgttatgtatatatatatttatgtatgtatatgtatatgtatatgtatat 1460
QY 1140 GGAGTGTGGCGCAGCTGTGATCACAGTACTCGGAGGCTGAGGACAGAGATCGCT 1199
|||||
Db 1461 gtagtgggcaagcgcgtgtgatacagctactcggaggtcgtgagcagagaaatcgct 1520
QY 1200 TGAGCCGCTGAAGTGCAGGCTGCTAGTGAGCCAGATGAGCCACTGCATTTCCAGCTGG 1259
Db 1521 tgagccgtgaaagtgcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1580
QY 1260 CGAAGAGAAAGACGCTGTCTCAAAACAACAAACAAGACTACTTTCAGACGTGTAG 1319
Db 1581 cgaagagaaagacgctgtctcaaaacaaacaaacaaagctactcttagcagtggttag 1640
QY 1320 AGTATCTCGCGCGGAAAGTGGGAAACGAGTGTGCACACAGAGTAGGCTATTTATATG 1379
Db 1641 agtatctcgcgggcggaagtgggaaacagtgctgcacacagagtaggcatactttatatg 1700
QY 1380 TTAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTTCTTGGCAGTGTATA 1439
Db 1701 ttaacagacactgaTaccagctaaagcggtgaacacatttactctctgagtggttta 1760
QY 1440 AAAGTATCTGTTTCTCATATTTTATTTTATTTTATTTTCTGGATCAAGCAACCTGAT 1499
Db 1761 aaagtatctgtcttctctcata tggtttatctttaaatttttcttgatcaagaacactgat 1820
QY 1500 CTTTTTCTCATAACTTGCAGCCAGCCGCTGACAGCAAAACCGCAGAAAGCTCGCCGAC 1559
Db 1821 ctttttctcataacttgccgacgacccgtgacagcaaaacccgscagaagctcgcgac 1880
QY 1560 CTCACACCCGAGTCTGAGGTAGTGGCCCGGACTACATTTCCAGAGGACTTCGCGG 1619
Db 1881 ctccacacccagtgtaggttagtgcccccggactacattttccagaaggcaactgcggg 1940
QY 1620 CACGCTTCTGCTGCTCGGCTGAAGGGAAGGGCCAATCCTGTAGTATCTCGGGAAGGA 1679
Db 1941 cagctcttcgctcgttggcggcctgaaggaggggccaatccctgagtctctcgggaagga 2000
QY 1680 GGTGTCCGAGCCGCGACCTAGAGATCCAGAGCCACAGCGACGCGCCCGGCCGCC 1739
Db 2001 ggtgtccggagcgcgacctagagatcccagaagccacagcgagcgccgcccgcgc 2060

QY 1740 ACTATTTCCAGCTCTGGGGCAGGGTGGGCCCCAGACTCCACATTCGCCGGGTAGTG 1799
|||||
Db 2061 actatttccagctctgcggcgaggggtggccagactccacttcccgcggttagtg 2120
QY 1800 CGACCTTAGGGCGGACTTATGTCCAGAGGCTTCGGGGCGGCTGCGCGCGGTGCC 1859
|||||
Db 2121 cgacctagggcgggacttcatgtccccagcagctcccgcgcggtgcgcgcggtgccc 2180
QY 1860 TAGTGTGGATGTAAAGCGGAGGTGGCGAGGGGACCGAGGCCAGGACTCTCTCTGGG 1919
Db 2181 tagtgtggagtgaagcgcgagg tggcgagggggagccagggcagactctcttggg 2240
QY 1920 GTTTGGGGCTTGACTGGGTGCGCTTTCTGACAGACTTTACAGCCCCCGGGGCACAG 1979
Db 2241 gtttggggcttgacctgggtgcgctttctggacagactttacag-ccccgggggacag 2299
QY 1980 TCGTAGAGAGGGGGGGGGGCCCATTTGGGGCTCTCATTTGGGGTGTTCGGGGCGCACCC 2039
Db 2300 tcttagagagggggcgggcgccatttgggctctctatgggtgcgttggggcgacccc 2359
QY 2040 CATCGGTACCGGGCTCCCGGAATTTGGGGACAAAAGGCTCTGCAGTCTCGCTGA 2099
Db 2360 catcggg taccggggcgctcccggaatctgtgggggacaaaaggctctgcag tctcggtcga 2419
QY 2100 GGGTCTCACCCACAAAAGAGGGGAGCGGTGAGCAGAGGCTGAAGAGGTTGGGCAAGC 2159
Db 2420 ggggtctcaccgacaaaagaggggaagccggtgagcagaggtggaaggggtggggagagc 2479
QY 2160 AGGGAGCTGTGCTGTGCTCGGGGGGTGGAACTTACGCGGACCTTGGAGGAGGCTCC 2219
Db 2480 agggagagctgtgcgtgtgcggggcggtggaaccttagcgagacctgggagagagctccc 2539
QY 2220 CCGGCCGAACCTGCCCGACCTCCCTCCCGGGCTTGCCTTGCAGCGCCGCCACCGC 2279
Db 2540 ccggcgaaactgcgccgaacctccctcccccggtctgccttgcagggcgccgacacgcgc 2599
QY 2280 CCGCGCCCATGGGGCTGAAGGCCGCCAGAACAGCTGTTCGCCGTGCGCTCCATCGACG 2339
Db 2600 ccgcgcccatggggtgagggcgcccgagagagcgtgttccccgctgcgtccatcgagc 2659
QY 2340 ACGTGTGCGGCTGTTGCTGCCGAGCTGGCCGAGAGAGGAGCGGACCTGTGCTCTCTT 2399
Db 2660 acgtggtgcgctgtgtgctccgagctggcgagagagccgagacctgtgtctccctt 2719
QY 2400 CTTTGGTCTGCTGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTCATCTCCATCAACG 2459
Db 2720 ccttgggtgctggctgtgagcaatttcttgcgtgtcaaccgctcatctccatccacag 2779
QY 2460 TTCCGAGCTCAGCTTCAGCCCGCCCGCCCGCCCGAGCGGCTGGGGCTTCACCTACT 2519
Db 2780 ttccgagctcaccttccagcccgccccgaccccgctggcggtcctcaactact 2839
QY 2520 TTCCGCTGGCGACCTGTCTATCATCGCGGCTCTATGCGCGCTTACCGCCAGATCC 2579
Db 2840 ttccggtggcgacgtgtctatca tgcgcgctctctatgcgcgcttcccccagacc 2899
QY 2580 GAGCGCGCTGACCTGTCTCTATCTTCGAGAAAGGGGTGTCTCCAGCGGTGAGCTGG 2639
Db 2900 gggcgcgctgcacctgtccctctatctctgcagaaggggtgtctccagcggtgagctgg 2959
QY 2640 TGAAGAAGGCTCCGATGTATATGGAACAGCCTCAGCGGCTCTACTTCAAGGATCGGG 2699
Db 2960 tgaagaaggtctccgagtcatatggaacagcctcagcgctcctacttcaagggtcggg 3019
QY 2700 CCCACATCCAGTCCCTTTCAGTTCATCAGAGTTGGAGCCAGTAGTGGGAATCTTA 2759
Db 3020 cccacatccagtccctcttcagcttca tcaaggttggagccagtaggtgggaaatctta 3079
QY 2760 TCCATGACCCACTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAG 2819
Db 3080 tccatgacccacttcttcaaaacccctccatggtttcacagaaccttcttaagaactgttag 3139
QY 2820 CTTGTGAGGTTCGGCAGGTGTATTATTTCTCTTTCAGTTCGGGAAACTGAAGCCACAG 2879

|||||
Db 3140 ccttggaggttcgaggtgtatttctctcttgcagttggaaactgaagcccgag 3199
QY 2880 AGGGGAAATGATATGCAAAAGTCACACACGGCATGCGAGGGCTGGAAAGTGAAGCTGATC 2939
Db 3200 aggggaaaatgatatgcgaagtccacacacgcatggaaggtggaagtggaagcctgac 3259
QY 2940 ACTTGCTCCAAATCATCAACCTCACTCTGCCCCCTCAGCACCCCCACCCCTTGCCACTG 2999
Db 3260 acttgctccaaatcatcaacctcaactctgccccctcagcaccaccaccttgccactg 3319
QY 3000 AACAGCTACAGGAGTCTAAGCATGAGACACAGAGGGCGGCACAGATTTAGGGGGCAAG 3059
Db 3320 aacagctacagaggtctaaagcatgagacacagagggcgacagatttagggggcaag 3379
QY 3060 AAGATGAATTTGGCTGCACTTTGAGCGAGTTAAACAAAATAATGGCTATCAAGATTTTT 3119
Db 3380 aagatgaattgggctgcatattgagcgagctaaacaaaaaaatggctatgaagattttt 3439
QY 3120 TTTTTTTTTTTTGGAGACAGGCTCTCACTCTGTCCCCAGGCTGGAGTGCAGTGGTGT 3179
Db 3440 tttttttttttttgagacaggtctcaactctgtccccagcgtgagtgcaagtgtgt 3499
QY 3180 GATCATGGCTCACTGACGCTCACTCTCCCTGGGCTCAGAGATCTCCCAACCTCAGCCTC 3239
Db 3500 gatcatggctcaactgacgctcagtcctcctgggtcagagatcctccaaacctcagcctc 3559
QY 3240 CTCAGTAGCTGACAGTACAGGCGATGACCGTGGTGGTAAATTTTGTATTTT 3299
Db 3560 ctgagtagctgagagtagagggcatgacaggtggctggtgttaatttttttttttt 3619
QY 3300 GTAGAGATGGTGTCTCACTATGTGGCCAGACTGGTCTTGAACCTTTGGGCTCAAGTGAT 3359
Db 3620 gtagagatggtgtctcaactatggtggccagactggtctgaactcttgggtcaagtgtat 3679
QY 3360 CTCGCCGCTCAGTCTCCCAATGCTGGGATACAGGTGTGAGCCACCGCAACTGTGTGC 3419
Db 3680 ctgcgcgctcagctcccaaatgctgggattacaggtgtgagccaccgcgaactggtggtgc 3479
QY 3420 CTATGAAATTTTTTTTTTTTTCAGAGCGCGTCTCACTCTGCGCCAGCGTGGAGTGC 3479
Db 3740 ctatgaaaatttttttttttttttcagacgggtctcaactctgtcgccaggtggaggtgc 3799
QY 3480 AGTGGTCGAATCTCGGCTCACTGCAAGCTGTGCTCTCTGCTTCATGCGCAATCTCCTGCC 3539
Db 3800 agtgggtgcaatctcggctcaactgcaactgctcctcctgtttcaatgccaattcctgc 3599
QY 3540 TCTGCTCAGCTCCTGAGTACTGGGACTACAGGAGCTGCCACCATGCCCTGGGTAAAT 3599
Db 3860 tctgctcagctcctgagtagctggactagagagagcctgccaacctgctggctaat 3919
QY 3600 TTTTTTTTGGATTTTAGTAGACAGAGGTTTCCACCATGTTACCGAGATGGTCTGATC 3659
Db 3920 ttttttttggatttttagtagaacagaggtttcacactgttagccaggtgggtctcgatc 3979
QY 3660 TCTGAGCTCGTATCGGCCGCTGGCTCCCAAGTCTGGGATACAGGCGTGAGC 3719
Db 3980 tctgacctgtagtcgcccgccttggctcccaagtgtcggattacagggcgtgagc 4039
QY 3720 CACCGACCTGGTCAAAATGTTGAGACAGAGAGGGGTTGACCTCAAAAGGCTTAAG 3779
Db 4040 caccgacacctggtaaaaaatgtttgagacagagaggggttgacctcaaaaggcttaag 4099
QY 3780 ASTCAGGGCTTGCAAGAGCTTTGCACCAAGCCGGTTGACTGGCAATCCCATCTGGTG 3839
Db 4100 agtcagggcttgcaagagagctttgcaccaagcccggttgactggcaatcccaacctgggtg 4159
QY 3840 TGCCATATTGAGAGGAATCAGAGGCTGCTCTCAGCTTAGCAGGAAAAGAGTGCAGAGA 3899
Db 4160 tgccatatgagaaggaatcagaggtgctttcagcttagcagggaaaaaagagtgacagaga 4219
QY 3900 TAAATGAGGGTTATTTGTTGGTGGGTATATAGCCAGAGAGTGTGGCCAGCGTCTGTTT 3959
|||||

Db 4220 taaatgagggttatttgggtggtatagccagagagtggttggccagcgtctctgtt 4279
QY 3960 TTGCCATTCTCTGTTTAACTAGTAGTCAGTAAATGAAATGCCCTAAATCCATCAAT 4019
Db 4280 ttgccattctctgttttaacctagtaagtgcagtaaaatggaaatccctaaatccatagaat 4339
QY 4020 ATATAATAGAGTTGCGAGAGAAAGACGAGTAGGGCCAAAGGCTGGGTACGTACAGGATA 4079
Db 4340 atataatagagttgcagagaaagacaggttagggccaaagcctgggtcagctacaggata 4399
QY 4080 TCCAGAAAGGTATCTTGTGGACATAGAGGGTGTAAACACAGGGAGAGAGTCTTTCAACACG 4139
Db 4400 tcagaaaagtatacttgttggacatagaggtgttaaacagggagagagttcttgaacacg 4459
QY 4140 TGGAGGGAAGGATGAGGATAGTGGCAGGAGATCTGAGGTTGGGTACACAGGCTTG 4199
Db 4460 tgggaaggaaggtatggagaggaatagtgggcagaggaatctgaggttggttcaacagcgttg 4519
QY 4200 GAAAGGAGTGGAGGAGTGTGGCCCATCACTACCTGG-CCCTTTTCCCCTATGTTAAAG 4258
Db 4520 gaaagggagtgggagggaggtgtggcccatcactacctggcccccttcccccatgttaag 4579
QY 4259 CACAGAGGACCTCTTTCATTACTTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 4318
Db 4580 cacagaggacctcttcttaacctcccccttcacagggcaccaaaattggacagctcgg 4639
QY 4319 TGTGGCTTTGCTGTGGTGGGCTGCCAGGCGCTGGGTCTCGGGATGTCCACCTCGC 4378
Db 4640 tgtggcccttgcgtgtgttggccctggccagggccctgggtctccggagtgtccaccctcgc 4699
QY 4379 CCTGCTGAGGATCATGCCCTGGGTAGTGTGGGCCCAATGGGAGCAGACAGCTGAGGT 4438
Db 4700 cctgtctgaggtatcatgctcgtggttagtgttggggcccaatggggagcagacagctgaggt 4759
QY 4439 CACTGGCAGCGGAAGGGCAACGAGGACCGCAGGGGCCACAGTCAATGCCGTGTGGC 4498
Db 4760 cactggcagcggaaggcaacgagggcagggcgagggcgacagagtcgaatgccgtgtggc 4819
QY 4499 TGAGCGGTTATTGCTCCCTCCCGCAGCTTGTCCCTTTCATCTACTGTAGTACGCCAAGCC 4558
Db 4820 tgaagcgggtatgttccctccccccagcctgttccccctcactactgttagtgcacagcc 4879
QY 4559 ACCCAAGGACTCCATTTTCTTGGGCCACACCCCTTTCCTCCCATCACCCACATAGG 4618
Db 4880 acccaagggactccatttcttgggccacaccccttcttcccatcaccaccaLagg 4939
QY 4619 AAGGAAAGACAGAGACCCCTTTTCTGCTGTCTATTCCTGAAGCAGGACAGGCTGG 4678
Db 4940 aagggaagacagagagcccttttctgggtgtctatctcctgaagcaggggtgg 4999
QY 4679 GCCATCATGAGACATAATGATCTCATCCCCCTTAAGAGCTGGCTGACCTGAAAGGATC 4738
Db 5000 gccatcatgagacataatgatctcatcccccttaagagctggctgLaacctgaaagatc 5059
QY 4739 ATACATCGCTGTGACCGCAAGATGGAGTGGCTTTCATGTGTGTGTCCTCAACCCCTTC 4798
Db 5060 atacatgcgtgtgaccgcaagatggaggtgggttcatgtgtgtgtccatcaacccttc 5119
QY 4799 CATTGACCTGCACACCGACTCGGTGGAGTCTGTGCAGCTGCAGCAGGTGAGGGCTGAGCC 4858
Db 5120 cattgaactgcacacgactcgtggagcttctgcagctgcagcaggtgagggctgagcc 5179
QY 4859 AATGGGGCAGGACTGGGCTAGGCCAGACTTGTACTGTGTGGGACCTGGCGCAGGGCAC 4918
Db 5180 aa19999caggaactgggctagggccagactgactgtgtgtgggacctgggaggggacac 5239
QY 4919 TTTCCCTTCTGAGCTTCAGCTTCCCTCCTCGGAAAATGGGTTAGTAATTTCTTGGGCTG 4978
Db 5240 ttccccctctgacttgcagctccccctcctcctggaaaaatgggttagLaa1tctcggcctg 5299
QY 4979 GCCTTTCCCGAGGGCTCTTGGGAGAGTAGAATTGAGATGTGAAATTTGCTTTCACCTCA 5038
Db 5300 gcctttccccagggctcttgggagagtagaattgagatgtgaaattgtcttgaactccatta 5359
|||||

QY 5039 AAGGCTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGGTTCTTGAC 5098
|||||
Db 5360 aaggctgagtcgccaaatttggcccttccacatggtgggtggttctgac 5419
|||||
QY 5099 CCCACCTCTGCCGATAGGCTAAGGACCGGTTCCTCCCTGTTCCTGGCTGATCAACT 5158
|||||
Db 5420 cccacctctgcccgataggttaaggaacccgttctccctccctgttccgtggctcataact 5479
|||||
QY 5159 CTCCTCTTGGCTCCTACAAAGCTGTCTGGCTGTCTATGACCTGGGACATCTGGAAGG 5218
|||||
Db 5480 ctctcccttcggctccatagaagctgcttggtgctctctatgacctgggacactctggaaagg 5539
|||||
QY 5219 TCAGTAGAGGAAGTGGCAGGCTGCGCTGTGTAGGCGGGGGGCTGGGTGGCAGCCTG 5278
|||||
Db 5540 tcagtagaggaagtggccaggctgctgctggtgaagcgcgggggtggtggcagcctg 5599
|||||
QY 5279 AATTATGATCTTCTTAGGTACCCCATGGCTTAGGGAACCTGGCAGATCTAGAGGAGC 5338
|||||
Db 5600 aattatgatcttctctaggttaccctatggccttagggaacctggcagatctagaggagc 5659
|||||
QY 5339 TGGAGCCACCCCTGGCGGGCCAGACCACCTCACCTCTACCAAGGTGGGGGCATCTA 5398
|||||
Db 5660 tggagccacccctggccggccagaccactcaacctctacacaaaggcgggggccatcta 5719
|||||
QY 5399 AGGAGGGTGCAAGAGGAGACCCTTAACAGTGGCTGAGGCGAGGGGCCCTCATCTGGGCAGA 5458
|||||
Db 5720 aggagggctgacagaaggagaccctaaacagtggctgagcgaggggccctcatctgggcaga 5779
|||||
QY 5459 TGAGAAGAGAACTTGTGTGTGGGGGTATCGGCCATCCAGTCTCAGTTGTGTCAACT 5518
|||||
Db 5780 tgaagaagaactttgtgtgtgggggtatcgccatccagctcactttgtgtcaact 5839
|||||
QY 5519 GTGTGCAAACTCAGTTCAAGTCTCAGGCTGTCTGAGGGGTGTCCAGGGTTCGCCAGCCTGGG 5578
|||||
Db 5840 gttgtcagaaatcagttcagttcagggctgtctgaggggggtccaggggttccccagcccggg 5899
|||||
QY 5579 AGTGGCAGGGGTGCAATTTGCCCTCAGCCCTCAGCCCTGCTTTCTGCCACTGCTACTGTCT 5638
|||||
Db 5900 agtggcaggggtgctgatttgcctcctcagcctgctcttctgccaetgcttactgtct 5959
|||||
QY 5639 TCCTGGAGTATTAACAGAGTCAAAATGTGGTAGGACCTGAGAGGGGTCTTCACTTGG 5698
|||||
Db 5960 tcttggaagtaaaacagaggttcaaatgtgtgtaggagcactgaagagggggtttcaacttgg 6019
|||||
QY 5699 TGGGTGTAGTGGGAGAGAGGCCATTTGGCTGGGCTTTGAAGTCTTTGGTGTGATGTAG 5758
|||||
Db 6020 tgggtgtaggtggggagagggccattgggctgggcttgaagcttcttggtagtgtag 6079
|||||
QY 5759 AAGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGGAGGGCAGGCCCCACCCCTGCACTC 5818
|||||
Db 6080 aagagtgtctgagaaagagaaggccctgagctcgagggcaggggcccccacccctgagctc 6139
|||||
QY 5819 TGCCCCAGGCTCAGCCAGCAGTCTCTGTAGACCCAGGAGGAGACCAGGTAGAAGGGCTG 5878
|||||
Db 6140 tgcctccagggctcagccagcagctctgtgtagccacaggaggagaccagggtagaagggtg 6199
|||||
QY 5879 GCAGCGAGTGGAGGTGGAGGTGGAGATGGAGGAGTCCCTGGGATCTTCTGTGGCCCC 5938
|||||
Db 6200 gcagcgagtggaggtggaggtggagatggagaggaactcccttgggatcttctctgtggcccc 6259
|||||
QY 5939 TTCTGGGTGTGCCCTGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGCTCCCTCCCTGA 5998
|||||
Db 6260 tctctgggtgtgcccttggggcatttgtgcagcagggcagctgggggtgctccctcctga 6319
|||||
QY 5999 GGATCTCTGCTCAGCTTCATCCAGGGCATTTGCTCAGCCAAAGACCTACTATCGGGATG 6058
|||||
Db 6320 ggaatcctgtcctcaactccatccagggcattgtcctcagccaaagacctactatcgggatg 6379
|||||
QY 6059 AACACATCTACCCCTTACATGTACTGTGGCTGTACCTGTCCGAAACCGCAATGTGGGG 6118
|||||
Db 6380 aacacatctacccctaaatgttaccttggctgtaccactgtcgcaaacccgcaatgtgcgg 6439

QY 6119 AAGCCCTGACGAGCCTGGCGGACACAGCCACTGTATCCAGGAGTGAAGATCCCCCTACT 6178
|||||
Db 6440 aagccctgcagccctggggagacagccactgtcatccagagtgagatccccctact 6499
|||||
QY 6179 AGGCTGTGACGCTGTCTTCTTCTCCCTTCCATCATGAGTTTCAACACACCTCTGTCAGGAC 6238
|||||
Db 6500 agggcctgcagccctgtccttctccctccatcagtttccaaacccacccctcgtccaggac 6559
|||||
QY 6239 TGAGGCTGTGGCTCCACAGCCCATCCCTTTTCCATCCAGTCCCTAGGCAGCAAGGCCACC 6298
|||||
Db 6560 tgaaggcttggctcccaagccccatcccccttccatccagtcctcagcagcaagccacc 6619
|||||
QY 6299 ATTACCAGGAGGTAGGACCCCTGATTAAAGTGTACATCTTTCCTCCCTCCCTCTCC 6358
|||||
Db 6620 ataccagaggttagggacccctgatlaaggtgtaacatcttccctccctccccctctcc 6679
|||||
QY 6359 TCCTAATTTTTTTTTTCTCAGAACAGTCTCAANTCTCAANTCTTCAANTCTTAAACCACCATCATCCA 6418
|||||
Db 6680 tcttaatttttttctcagaacagttctcaaatctcaatgtttaaccacacatcata 6739
|||||
QY 6419 GCAGTGGGACTTCCACCCCTGGCCCCATGCCCTCTCTCATTTCTTGTCTTCTCTCTG 6478
|||||
Db 6740 gcagtgggacttccacccctggccccctgccccctccctcattcttcttctctctg 6799
|||||
QY 6479 GGCTGACCCAGACAGCATCATTTTTTGCAGTGAAGAGCCCACTACTCCCCAGCCCTTGGG 6538
|||||
Db 6800 ggcagaccagacagcatcattttgcagtaggagagcccaactctccccagccccctggg 6859
|||||
QY 6539 GGCTTCATCCCCCGCCAGGTCCCTGGGGCTACCCCGCATGTGTGAGACCCCTTTCAGACCCCT 6598
|||||
Db 6860 ggcctcatcccccccgaggtccctgggggtacccccgataggtagagcccccttcagacccct 6919
|||||
QY 6599 ACAGAGACCCCACTGCTCTCACAGTACAACTACTTGCCTGGGAAGAGAGGAGTCTACAA 6658
|||||
Db 6920 acagagaccacactgctctcagactcaactactgcccggagagagagagagagactcaaa 6979
|||||
QY 6659 GGAATCTTTGAAGTAGCAATGATATCCCAACTGCTGAAGAGGAGGAGGAGGAGGAGT 6718
|||||
Db 6980 ggaattctttgaatagccaatgatgtaatcccccaactgctgaaggagggcagcagctt 7039
|||||
QY 6719 GCTGAGGCGGGCGGAGGAGCGGGGAGCAAGCCAGGTGAAGGCTGAGGCTGAGCTCCAG 6778
|||||
Db 7040 gctggagggcgagagagagcgggggggagcaaaagcaggtgaaaaggtggaggtccag 7099
|||||
QY 6779 CCTGTGTCCAGCCCTCCACCTTGGACAGGGCTCCCTTCCACAGGGCCATGGGGGTGCAAG 6838
|||||
Db 7100 cctgtgtccagcctcccaacttggacaggggtcccttccaca-ggccaatgggggtctgcatg 7158
|||||
QY 6839 TACGGGATTAGGATGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6898
|||||
Db 7159 tacgggattagggatggcagggagaaagggtggcctgagcagacagctatgttcccttttg 7218
|||||
QY 6899 CTATACTGAGGTCTGGGCCACCTTGGAGGGGACTGAAGGTATTTAGAGGTCTTCTAC 6958
|||||
Db 7219 ctaaaactgaggtccctggggccacgttggagcggagctgaaagggtatttttagagggttctcac 7278
|||||
QY 6959 CCTGTGCTTCAAGTTTATGCGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGAGGAGT 7018
|||||
Db 7279 cctgtgcttcaagtttcatggcagactccctccctcagctgaggggtgagaggtagggat 7338
|||||
QY 7019 GGTACGCTCTGGCTATGAGTTGGCTTTATAAAGAGAAAGAGGTTCTTAAGAAATGTTCACAA 7078
|||||
Db 7339 ggtacgtccctggctatggatttggcttataaaaggaaagaggttctcaagaatgttcccaa 7398
|||||
QY 7079 CCTATGCTTACTTTTCTGGAGCCAGGGGTCTTTTGCCTAGGTGGGGGCCCTGGGCTTGC 7138
|||||
Db 7399 cctatgcttaccttcttggagccaggggtcttgcctaggtggggggcctggcctgtgc 7458
|||||
QY 7139 CCTGTCTTAAGGGGTGAGTAAAGAGACTGATCTGTGCCCTCCCTCCCTCTGCTGCTCAGGGC 7198
|||||
Db 7459 cctctgctaagggtgtgagtaagagactgactgtgcctcccttccccctcagtcacagggc 7518
|||||
QY 7199 ACCCAGAGCCCAAGGTTCCGCCCTCCAGGAGCCCTGAGTGTCTTTCGCCCACTCCGATTC 7258

Db 7519 accagagccaaaggttcgccctccaggaccctgagtgttcgccacacctgctgcgatic 7578
QY TACGACGGCATCTGCAAAATGGAGGAGGCGAGTCCACACCGCTGTGCTGCACGTGGGCTGG 7318
Db 7579 tacagacggcatctgcaaaTgggagggggcagTcccacacgctgctgctgacgTgggctgg 7638
QY GCCACCTTTCTTGTGAGTCCCTAGGCCCTTTTGTAGGGACAGGTGAGGACAGCTGCACA 7378
Db 7639 gccacctttctgtgcagTcccctagccgttttgaggagacagtgaggaacagctgcaca 7698
QY GAGGCTGGGCACTACAGTGTGTGACAGACGACCGCGGTGTGACAGCTTTCTTGCCCA 7438
Db 7699 gaggtctggcactacagtggtgacagcagccagggctgttcagactttcttgccca 7758
QY GGGGACGATCTGCCATCCCTTCGGTCCGATGGAGTGGAGCCCTGGTGGGATG 7498
Db 7759 ggggagcatctgcccacccccctcggtgcccagtgagacccccctgggtgggagTg 7818
QY GGATGCCACAGACAGGGTCTCTGGAGTTCCAGGCACCTGGCGGCAACTTGTCTCACCTT 7558
Db 7819 ggaTggcagagcagggtcctgaggttcagccactTggcggcaacctTgtctcaactt 7878
QY GCTCTCCCACTGGCCCAAGGTGGCGAGAGGTGCGCATAGTGAGCCGAGAGCCGAGGC 7618
Db 7879 gctctcccactTggcccaggtTggcgagaggtTgcgcatagTgagccgagagccgagc 7938
QY GCGCAGGCGGAGGAGCCCTGGGGCAGAGACGCCGGGAGGCCGCGGGGCCACG 7678
Db 7939 gggcgaggccgagggccgtTggggcgaggaagcccgaggccggcgggcccaTg 7998
QY GCGGAGTCCAAAGCAGAGGAGCCCGCCGCAAGAAAGCAGCAGTGGCAAGGGCCT 7738
Db 7999 gcggaTccaagcagagagagcccgcccgcccaagaagcagcaactggacaaagggcct 8058
QY GGGCACCGGCGAGGTGTGAGTGTGAGGACCCCGCGGAAGCTCTGAGGACTGTCTGCTGG 7798
Db 8059 gggcacggccagggTgcagTgcaggaccccccggaagcctctctggactgtcgctgg 8118
QY CACAGCCCGAGCCCTGAAGTGGCAGCAGCGTCAAGTGGCAGCACCCGCGACATCACC 7858
Db 8119 cacagcccgaggccctTgaagTggcgagcagcgtcagTggcagcaccgccagcatccc 8178
QY ACGCCGGAGGGTCCAGTGTCTCACTTTCAGAGTGAAGAAGTGAAGGATGAAGGAGCT 7918
Db 8179 accgcgagggTccagTgctcaacttcacagatgagaaagtgaaagcagagct 8238
QY GCTGTGGCCACCAAGATCACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCCGA 7978
Db 8239 gctggTggccaccaagatcaactcgagcgccaTcaagctTgcaactcacggcacagTcgca 8298
QY AGTGACAGTGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTCTTCTCTCAA 8038
Db 8299 agTgcagaTgaagaagcagaagTgtcccacccccctagTgactcaactctgtcttctcTaa 8358
QY GCGGACGCCAAAGGCTCTGAACTACTGGGACTTTCGACCGCTTGTGGGACCCAGGC 8098
Db 8359 gggcagcgcaaaagcctctgaactactggggaacttcgaacgcttgtgggaacccagc 8418
QY TCGG-CCTTAGTCCCCCACTGTGAGCCCATGTTCTGCCCCCAGGCCAACGAAGGGGACAGC 8157
Db 8419 tcgcctctagTcccccaactTgagcccatgtTctgccccagcccaaggggacagc 8478
QY CTCACCTCTACCCAAACCTAGTTCCTCCGTCCTCCGAGTACGTCTCTATCAACCCACGA 8217
Db 8479 ctcaactctaccaaaccttaggttccccggttccccgagtagctgtatcaaacccacga 8538
QY TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTGGTGTAGAAATAGTCTCTTCTCC 8277
Db 8539 ttttctcaagctcagaacccagggctctgccccagTcgTtagaataaggtctctctccc 8598
QY CAGAATCCAGCGCGGCAATGGAACCTCACGCTGGGTCTTAATACCAAGTCTTTAAAG 8337

Db 8599 cagaatcccagcgcccaatggaaacctcacgctgggtcctaataaccagttctttaaag 8658
QY CCCAGCCCTAGAAACCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCT 8397
Db 8659 ccagccccctagaaaccccaagctctctctcggaacccgctcacctagagccagccaagc 8718
QY TACTCAGGCTCCTCCAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCAAAGGAGCAC 8457
Db 8719 tactcagggctctctccagctTgtagagctgaggtttcaccccttaacccaa-ggagcac 8777
QY AGTCCCACCTCCAGCCCGGGAGCCTAGGACCACTAGCCCTAGGAGTATATTTCCCG 8517
Db 8778 aggtcccaactccagccc-gggagcctaggaaccaactcagccccctagagtatattccgc 8836
QY ACTTCAGAATTCATATCTTGGCAATCCCAAGCTCCCTGCCCAAAATAACTTTCAGTCTGC 8577
Db 8837 acttcagaattccatactctggaatcccaagctccccctgccccaaataactctcgtctgc 8896
QY TTCAGAATTTGGAATTCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8637
Db 8897 -tcagaatttggaatacctagtttctctctctctctctctctctctctctctctct 8955
QY TCGCCCCCAGCCTATGAGCATCCTGAGCCCCGCCCTCTTCTTGAGCAAACTGGCCCCGG 8697
Db 8956 tcgcgcccgagcctaTgagcatcctgagccccgcccctctctctgacgaaactggccccg 9015
QY ATCAGAGCAGGACCTCCCTCCGACCTCTGGAACCTCCAGAGTCCAGCCCATCTCG 8757
Db 9016 atcagagcaggacctccccctccagccctctggaaacctccagaggtccagcccatctgc 9075
QY GAGCATCCCGGAGGAATCTGCAGAGGGTTAGAGTGGGTGAGCAAGAGCTGATCTCTT 8817
Db 9076 gagcatcccgaggaatactctgcaga-gggtagagtgggTgacaagagcctgatctctt 9134
QY CTGTTTTGTACATAGATTATTTTTCAGTTCCCAAGAAAGATGAATACATTGTTTAAAA 8877
Db 9135 cctgtttgtacatagatttttttcagttccaagaagaatgaatcacatctgttataaa 9194
QY AAAATAAAGCCCAAGTGCATCTTATCTGGAATTTGGGATGGGCGGAGTGGAG 8937
Db 9195 aaataaagcgaagtcctatctatctggaaatctgggagatggggcgggagTggag 9254
QY CGCCCCCTTCTCTCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8997
Db 9255 cgccccct 9314
QY CGAGCAGAGTGGCGGGAAGACAGTAGGCTCGCCCCCGGCTTGGGGCTTCCCCCGCG 9057
Db 9315 cgagcgacagTggcgggagagcagtaggctccgcgcgcttggggcttcccccgcg 9374
QY CGCCCGGAGGCGCTCCCGGGCGCTCTCTCCCGGACTGGCGGTGGGCTATCCNGGC 9117
Db 9375 cgcccgagggccgctcccgcgccctctcccgactggcggtggggatcccc-ggc 9433
QY GCGGCCCGCCCGCGGCTTCAGCCCGCCCGCGGCTTCAGAGCCACGCGGCGCCCGCC 9177
Db 9434 gggcgcccgccccgggttcagcccccgccccggcttcagagccagcgggcgccgc 9493
QY 9178 CCG 9180
Db 9494 ccg 9496

RESULT 5
US-09-760-475-3945/c
; Sequence 3945, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper

Query Match	87.9%	Score 8067.2	DB 6	Length 8185
Best Local Similarity	99.8%	Pred. No. 3.1e-150		
Matches 8179	Conservative	0	Mismatches	Indels
QY	699	GGGTAAGGTGGTTCACGCCCTGTAATTTTAGCACATTTGGGAGCGCTGAGCGGCGCAGATCAC	758	
DB	8185	GGGTAAGGTGGTTCACGCCCTGTAATTTTAGCACATTTGGGAGCGCTGAGCGGCGCAGATCAC	8126	
QY	759	TTGAGGTTAGGGTTTCGAGACACAGCTCTGGCCAAACATGGTGAACCCCATCTCTACTTTAAA	818	
DB	8125	TTGAGGTTAGGGTTTCGAGACACAGCTCTGGCCAAACATGGTGAACCCCATCTCTACTTTAAA	8066	
QY	819	ATACAAAAGTTTCTCGGTGGTGGTGGCGGACGCTATATCCACAGCTACTTGGGACATTT	878	
DB	8065	ATACAAAAGTTTCTCGGTGGTGGTGGCGGACGCTATATCCACAGCTACTTGGGACATTT	8006	
QY	879	TTTTTAAGACGGAATCTCACCTCTGTGGCCAGGCTGGAGTGCAGTGGCAAGATTCCTGGCT	938	
DB	8005	TTTTTAAGACGGAATCTCACCTCTGTGGCCAGGCTGGAGTGCAGTGGCAAGATTCCTGGCT	7945	
QY	939	CACCTGAAGCCCTCCGCCCTCCACAGGTTCAAGGGGATTCGCCGCGCTCAGCCCTCCCAAGTAG	998	
DB	7945	CACCTGAAGCCCTCCGCCCTCCACAGGTTCAA - GGGATTCGCCGCGCTCAGCCCTCCCAAGTAG	7887	
QY	999	CTGGGAATCCCTGTCTCTGCAGAAAAAAGAAAAAAGAAAAAATATATATATATA	1058	
DB	7886	CTGGGAATCCCTGTCTCTGCAGAAAAAAGAAAAAAGAAAAAATATATATATATA	7827	
QY	1059	TATATATGTGTGTGTGTGTGTGTGTGTGTATATATATATATATATATATATATGC	1118	
DB	7826	TATATATGTGTGTGTGTGTGTGTGTGTGTATATATATATATATATATATATGC	7767	
QY	1119	ACATACACAAAAATTAGCGGGAGTGGTGCACGCGCTGTGATCACAGCTACTCTCGGA	1178	
DB	7766	ACATACACAAAAATTAGCGGGAGTGGTGCACGCGCTGTGATCACAGCTACTCTCGGA	7707	
QY	1179	GGCTGAGCAGAGAAATCGCTTGACCCCGTGAAGTCGAGGCTGCAGTGCAGCCAGATGA	1238	
DB	7706	GGCTGAGCAGAGAAATCGCTTGACCCCGTGAAGTCGAGGCTGCAGTGCAGCCAGATGA	7647	
QY	1239	GCCACTGCATTCCAGCCCTGGCGGAAGAGAAGACCGTGTCTCAAAACAAACAAACAAA	1298	
DB	7646	GCCACTGCATTCCAGCCCTGGCGGAAGAGAAGACCGTGTCTCAAAACAAACAAACAAA	7587	
QY	1299	GCTACTCTTAGCAGCTGTAGAGTATCTCGCGGGCGGAAGTGGAAACGAGTGTGCACA	1358	
DB	7586	GCTACTCTTAGCAGCTGTAGAGTATCTCGCGGGCGGAAGTGGAAACGAGTGTGCACA	7527	
QY	1359	CAGAGTAGGCATCTTTATATGTTTAACAGACACTGATACCCAGCTAAAGCGGCTGAACACA	1418	
DB	7526	CAGAGTAGGCATCTTTATATGTTTAACAGACACTGATACCCAGCTAAAGCGGCTGAACACA	7467	
QY	1419	TTTACTCTCTGGCAGTGTTTAAAGTATCTGTTTTTCTCATATGTTTTTATTTTAAATTTT	1478	
DB	7466	TTTACTCTCTGGCAGTGTTTAAAGTATCTGTTTTTCTCATATGTTTTTATTTTAAATTTT	7407	
QY	1479	TTCTGGATCAAGCAACCTGATCTTTTTCTCATTAACCTTGGCGACCGACCGCTGCACGAAA	1538	
DB	7406	TTCTGGATCAAGCAACCTGATCTTTTTCTCATTAACCTTGGCGACCGACCGCTGCACGAAA	7347	
QY	1539	AACGGCAGAGCTCGGGCAGCCTCCACCCCGAGTCTCGAGGTAGTGCCTCCCGGACTACA	1598	
DB	7346	AACGGCAGAGCTCGGGCAGCCTCCACCCCGAGTCTCGAGGTAGTGCCTCCCGGACTACA	7287	

Db	5127	GCTTGACCTCAAAAGGCTTAAGAGTCAGGGCTTGCAAAAGAGCTTGTCACCAAGCTTCGGTT	5066
Qy	3818	GACTGGCAATCCCATCCTGGTGTGCCATATTGAGAAGGAATCAGAGGCTGCTTCTCAGCT	3877
Db	5067	GACTGGCAATCCCATCCTGGTGTGCCATATTGAGAAGGAATCAGAGGCTGCTTCTCAGCT	5008
Qy	3878	TAGACGAAAAGAGTCGAGAGATAAATAGAGGTTATTGTTGGTGGGTGTATAGCCACAG	3937
Db	5007	TAGCAGAAAAGAGTCAGAGATAAATAGAGGTTATTGTTGGTGGGTGTATAGCCACAG	4948
Qy	3938	AGTGTGGCCAGGCTCCTGTTTGGCCATTCCTGTTTAACTAGTAGTCGCTAAAT	3997
Db	4947	AGTGTGGCCAGGCTCCTGTTTGGCCATTCCTGTTTAACTAGTAGTCGCTAAAT	4888
Qy	3998	GGATCCCTAAATCCATAGAATATATAATAGAGTTGCAGAGAAAGACAGGCTAGGGCAA	4057
Db	4887	GGATCCCTAAATCCATAGAAATATATAATAGAGTTGCAGAGAAAGACAGGCTAGGGCAA	4828
Qy	4058	AGGCTGGGTCAGCTACAGGATATCCAGAAAGATATCTTGTGGACATAGAGGGTATAAC	4117
Db	4827	AGGCTGGGTCAGCTACAGGATATCCAGAAAGATATCTTGTGGACATAGAGGGTATAAC	4768
Qy	4118	AGGAGAGAGCTCTTTGAACACGTGGGAGGAAGGATGGAGGGATAGTGGCAGAGAAAT	4177
Db	4767	AGGAGAGAGCTCTTTGAACACGTGGGAGGAAGGATGGAGGGATAGTGGCAGAGAAAT	4708
Qy	4178	CTGAGGTTGGGTCACAGGCTTGGAAGGAGGTGGAGGAGGTGGGCCCATCACTACCTG	4237
Db	4707	CTGAGGTTGGGTCACAGGCTTGGAAGGAGGTGGAGGAGGTGGGCCCATCACTACCTG	4648
Qy	4238	G-CCCCCTTCCCCATGTTAAAGCACAGAGGACCCCTCTTTCATTACCTCCCCCTTCCACAG	4296
Db	4647	GCCCCCTTCCCCATGTTAAAGCACAGAGGACCCCTCTTTCATTACCTCCCCCTTCCACAG	4588
Qy	4297	GCACCAATTTGGACAGCTCGGTGTGGCCCTTTGCTGTGGTTGGGGCTCCGACGGCCCTGG	4356
Db	4587	GCACCAATTTGGACAGCTCGGTGTGGCCCTTTGCTGTGGTTGGGGCTCCGACGGCCCTGG	4528
Qy	4357	GTCTCCGGGATGCCACCTCGCCCTGTCTGAGGATCATGSCCTGGGTAGTGTGGGGCCCA	4416
Db	4527	GTCTCCAGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGGTAGTGTGGGGCCCA	4468
Qy	4417	ATGGGGAGCAGACGCTGAGGTCACCTGGCAGCGCAAGGGCAACGAGGACCGCAGGGGCC	4476
Db	4467	ATGGGGAGCAGACGCTGAGGTCACCTGGCAGCGCAAGGGCAACGAGGACCGCAGGGGCC	4408
Qy	4477	AGACAGTCAATGCCGGTGTGGCTGAGCGGGTATTGTTCCCTCCCCCAGGCTTGTCCCT	4536
Db	4407	AGACAGTCAATGCCGGTGTGGCTGAGCGGGTATTGTTCCCTCCCCCAGGCTTGTCCCT	4348
Qy	4537	TCATACTGATAGTACCCACGCCACCAAGGACTCCATTTCCTGGGCCACACCCCTTC	4596
Db	4347	TCATACTGATAGTACCCACGCCACCAAGGACTCCATTTCCTGGGCCACACCCCTTC	4288
Qy	4597	TTCCCATCACCCACCATAGGAAGGAAGACAGAAAGAGGCCCCCTTTTCTCGGCTGTCAAT	4656
Db	4287	TTCCCATCACCCACCATAGGAAGGAAGACAGAAAGAGGCCCCCTTTTCTCGGCTGTCAAT	4228
Qy	4657	CCCTGAAGCAGGACAGGGTGGCCCATCATGAGACATAATGATCTCATCCCCCCTAAGA	4716
Db	4227	CCCTGAAGCAGGACAGGGTGGCCCATCATGAGACATAATGATCTCATCCCCCCTAAGA	4168
Qy	4717	GCTGGCTGTACCTGAAGGATCATACATCGCTGTGACCGCAAGATGAGGTGGCGTCA	4776
Db	4167	GCTGGCTGTACCTGAAGGATCATACATCGCTGTGACCGCAAGATGAGGTGGCGTCA	4108
Qy	4777	TGCTGTGTGCATCAACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGC	4836
Db	4107	TGCTGTGTGCATCAACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGC	4048
Qy	4837	TGCAGAGGTTGAGGGCTGAGCCAAATGGGCGAGGACTTGGGCTAGGCCAGACTTGCCT	4896
Db	4047	TGCAGAGGTTGAGGGCTGAGCCAAATGGGCGAGGACTTGGGCTAGGCCAGACTTGCCTGCT	3988

Qy	4897	GTGGGACCCCTGGCGAGGGCACTTTCCCTTCTGAGCTTCAGCTTCCCTCCTCGGAAAAA	4956
Db	3987	GTGGGACCCCTGGCGAGGGCACTTTCCCTTCTGAGCTTCAGCTTCCCTCCTCGGAAAAA	3928
Qy	4957	TGGGTTAGTAATTCCTGGCCCTGGCCCTTTCCAGGGCTCTTGGGAGAGTAAATTTGAGATG	5016
Db	3927	TGGGTTAGTAATTCCTGGCCCTGGCCCTTTCCAGGGCTCTTGGGAGAGTAAATTTGAGATG	3868
Qy	5017	TGAAATTTGCTTTGACTCCCATTAAGAGGCTGGTCCACAGAAATTTTGGCCCTTCCACATGGTG	5076
Db	3867	TGAAATTTGCTTTGACTCCCATTAAGAGGCTGGTCCACAGAAATTTTGGCCCTTCCACATGGTG	3808
Qy	5077	GGTGGTCCCTGTGTGTTCTGACCCCACTCTGGCCCGATAGGCTAAGGACCCGTTCTCCT	5136
Db	3807	GGTGGTCCCTGTGTGTTCTGACCCCACTCTGGCCCGATAGGCTAAGGACCCGTTCTCCT	3748
Qy	5137	CCCTGTTCCGTGGCTCATAACTCTCTCTTGGCTCTCTAGAAAGTGTCTGGCTGCTCTA	5196
Db	3747	CCCTGTTCCGTGGCTCATAACTCTCTCTTGGCTCTCTAGAAAGTGTCTGGCTGCTCTA	3688
Qy	5197	TGACCTGGGACATCTGAAAGGCTCAGTAGAGGAAGTGGCCAGGCTGGCCTGTGTGAGGC	5256
Db	3687	TGACCTGGGACATCTGAAAGGCTCAGTAGAGGAAGTGGCCAGGCTGGCCTGTGTGAGGC	3628
Qy	5257	CGGGGGCTGGGTGGCAGCCTCAATTTATGATCTTCTTAGTTACCCCATGGCCTTAGGG	5316
Db	3627	CGGGGGCTGGGTGGCAGCCTCAATTTATGATCTTCTTAGTTACCCCATGGCCTTAGGG	3568
Qy	5317	AACCTGGCAGATCTAGAGAGCTGGAGGCCACCCCTGGCCGGCCAGACCACCTCACCCCTC	5376
Db	3567	AACCTGGCAGATCTAGAGAGCTGGAGGCCACCCCTGGCCGGCCAGACCACCTCACCCCTC	3508
Qy	5377	TACCACAAGTGGGGCATCTAAGCAGGGTGCAGAAAGGAGACCCCTAACAGTGGCTGAGG	5436
Db	3507	TACCACAAGTGGGGCATCTAAGCAGGGTGCAGAAAGGAGACCCCTAACAGTGGCTGAGG	3448
Qy	5437	CAGGGGCCCTCATCTGGGCAGATGAGAAGAACCTTTGTGTGTGGGGGGTATGCCCCAT	5496
Db	3447	CAGGGGCCCTCATCTGGGCAGATGAGAAGAACCTTTGTGTGTGGGGGGTATGCCCCAT	3388
Qy	5497	CCAGTCTCACTTTGTCTCAACTGTGTGCAGAAATCAGTTTCAGTCAGGCTGTCTTGAGGGGT	5556
Db	3387	CCAGTCTCACTTTGTCTCAACTGTGTGCAGAAATCAGTTTCAGTCAGGCTGTCTTGAGGGGT	3328
Qy	5557	GTCCAGGGTTCCCGAGCCTGGAGTGGCAGGGGTGCAATTTGTCTCCCTCAGCCCTGCCTT	5616
Db	3327	GTCCAGGGTTCCCGAGCCTGGAGTGGCAGGGGTGCAATTTGTCTCCCTCAGCCCTGCCTT	3268
Qy	5617	TTCTGCCACTGCTTACTGTCTTCTTGAGATATACACAGAGTCAAAATGTGGTAGGAGCAC	5676
Db	3267	TTCTGCCACTGCTTACTGTCTTCTTGAGATATACACAGAGTCAAAATGTGGTAGGAGCAC	3208
Qy	5677	TGAAGAGGGGTCTTCACATTTGTGGTGTGTAGTGGGGAGGAGGCCATTGGGCTGGGCTT	5736
Db	3207	TGAAGAGGGGTCTTCACATTTGTGGTGTGTAGTGGGGAGGAGGCCATTGGGCTGGGCTT	3148
Qy	5737	GAAGATCTTTGGTGATGTGTAGAAGAGTGTCTTGAAAGAGAAAGGGGCCCTGAGCTCGGAG	5796
Db	3147	GAAGATCTTTGGTGATGTGTAGAAGAGTGTCTTGAAAGAGAAAGGGGCCCTGAGCTCGGAG	3088
Qy	5797	GGCAGGCCCAACCCCTGCAGTCTGCCCAAGGCCCTCAGCCAGCTCTGTAGACCCAGGG	5856
Db	3087	GGCAGGCCCAACCCCTGCAGTCTGCCCAAGGCCCTCAGCCAGCTCTGTAGACCCAGGG	3028
Qy	5857	AGGAGACCAAGTAAAGGGCTGCACCGCAGTGGAGGCTGGGAGTGGAGTAGGAGACTC	5916
Db	3027	AGGAGACCAAGTAAAGGGCTGCACCGCAGTGGAGGCTGGGAGTGGAGTAGGAGACTC	2968
Qy	5917	CCTGGGATCTTCTGTGGCCCTCTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCACAGG	5976
Db	2967	CCTGGGATCTTCTGTGGCCCTCTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCACAGG	2908

Qy	5977	GCAGCTGGGGGCTGCCCTCCCTGAGGATCCCTCTGCCCTCACCTCCATCCAGGGGCACTTGGCTTCA	6036
Db	2907	GCAGCTGGGGCTGCCCTCCCTGAGGATCCCTCTGCCCTCACCTCCATCCAGGGCACTTGGCTTCA	2848
Qy	6037	GCCAAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACTGGCTGGCTGGCTACCC	6096
Db	2847	GCCAAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACTGGCTGGCTGGCTACCC	2788
Qy	6097	TGTCGCAACCGCAATGTGGGGGAAGCCCTGCAGGCGCTGGCGGGACACGGCCTACCTGTCATC	6156
Db	2787	TGTCGCAACCGCAATGTGGGGGAAGCCCTGCAGGCGCTGGCGGGACACGGCCTACCTGTCATC	2728
Qy	6157	CAGGAGTGAGGATCCGCCCTACTAGGCGCTGCAGCCTGTCTTCTTCTTCCCTCCATCACTAGT	6216
Db	2727	CAGGAGTGAGGATCCGCCCTACTAGGCGCTGCAGCCTGTCTTCTTCTTCCCTCCATCACTAGT	2668
Qy	6217	TCCAAACCACCTCGTCCAGGACTGAGGCGCTGGCTCCAGCGCCCATCCCTTTCATCCCA	6276
Db	2667	TCCAAACCACCTCGTCCAGGACTGAGGCGCTGGCTCCAGCGCCCATCCCTTTCATCCCA	2608
Qy	6277	GTCCCTAGGCAGCAAGGCCACCATTACCAGGAGGTAGGAGCCCTGATTAAAGTGTGTACA	6336
Db	2607	GTCCCTAGGCAGCAAGGCCACCATTACCAGGAGGTAGGAGCCCTGATTAAAGTGTGTACA	2548
Qy	6337	TCCTTTCCTCCCTCCCTCTCCTCTAAATTTTTTTTTTTTCTCAGAACAGTCTCAAAATCTCC	6396
Db	2547	TCCTTTCCTCCCTCCCTCTCCTCTAAATTTTTTTTTTTTCTCAGAACAGTCTCAAAATCTCC	2488
Qy	6397	AATGTTTTAACCCACCATCATCCAGACAGTGGGACTTCCACCTCGGGCCCATTCGCCGCCCTCC	6456
Db	2487	AATGTTTTAACCCACCATCATCCAGACAGTGGGACTTCCACCTCGGGCCCATTCGCCGCCCTCC	2428
Qy	6457	TCATCTCTGGCTTTCTTCCTCTGGGCTGACCCAGACAGACATCATTTTGCAGTGAGGACCCC	6516
Db	2427	TCATCTCTGGCTTTCTTCCTCTGGGCTGACCCAGACAGACATCATTTTGCAGTGAGGACCCC	2368
Qy	6517	ACCTACTCCCGCCAGCCCTGGGGCTTCCATCCCGCCGAGGTCCTCGGGGTACCGCCCGA	6576
Db	2367	ACCTACTCCCGCCAGCCCTGGGGCTTCCATCCCGCCGAGGTCCTCGGGGTACCGCCCGA	2308
Qy	6577	TGGTGAGACCCCTTACAGCCCTACAGAGACCCCACTGCTCTCACAGCTTACAACCTACTGCCC	6636
Db	2307	TGGTGAGACCCCTTACAGCCCTACAGAGACCCCACTGCTCTCACAGCTTACAACCTACTGCCC	2248
Qy	6637	GGGAAGACAGGAGATCTCAAGAGAGTCTTTGAAGTAGCCCAATCATGTCTATCCCAACC	6696
Db	2247	GGGAAGACAGGAGATCTCAAGAGAGTCTTTGAAGTAGCCCAATCATGTCTATCCCAACC	2188
Qy	6697	TGCTGAAGGAGGAGCCAGCTTCTTGAGGCGGGGAGGAGCGCGCGGGGAGCAAGGCC	6756
Db	2187	TGCTGAAGGAGGAGCCAGCTTCTTGAGGCGGGGAGGAGCGCGCGGGGAGCAAGGCC	2128
Qy	6757	AGGTGAAGGCTGGAGTCCAGCCTGTCTCAGCCTCCCACTCGACAGGCTTCCCTTCC	6816
Db	2127	AGGTGAAGGCTGGAGTCCAGCCTGTCTCAGCCTCCCACTCGACAGGCTTCCCTTCC	2068
Qy	6817	ACAGGGCCATGGGGCTGCATGTACGGGATTAGGGATGGCAGGAGAGCTTGGCCCTGAG	6876
Db	2067	ACA - GGGCCATGGGGGCTGCATGTACGGGATTAGGGATGGCAGGAGAGCTTGGCCCTGAG	2009
Qy	6877	CAGACAGCTATGTTCCCTTTTGCTATAACTGAGGTCCTGGGCGCCACGTTGACGGGACTG	6936
Db	2008	CAGACAGCTATGTTCCCTTTTGCTATAACTGAGGTCCTGGGCGCCACGTTGACGGGACTG	1949
Qy	6937	AAGGATTTTTAGAGGTTTACCCCTGTGCTTCAGTTTTCATGGCCAGACTTCCCTCCCTTCA	6986
Db	1948	AAGGATTTTTAGAGGTTTACCCCTGTGCTTCAGTTTTCATGGCCAGACTTCCCTCCCTTCA	1889
Qy	6997	GCTCAGGGGTGGAGGTAGGATGGTACGTCCTGGCTATGGATTGCTTTATAAAGGAAA	7056
Db	1888	GCTCAGGGGTGGAGGTAGGATGGTACGTCCTGGCTATGGATTGCTTTATAAAGGAAA	1829
Qy	7057	GAGCTTCTAAGATGTTCCCAACCTATGCTTACCTTTTCTGGAGGCAGGGGTCTTTTGCT	7116

Db 1828 GAGGTTCTAGAAATGTTCCAAACCTATGCTTACCTTTCTTGGAGCCAGGGGTCTTTGCCT 1769
QY 7117 AGTGGGGGCGCTGGCTGTGCGCTCTGCTAAGGGTGAGTAAAGACATGATCTGCGCC 7176
Db 1768 AGTGGGGGCGCTGGCTGTGCGCTCTGCTAAGGGTGAGTAAAGACATGATCTGCGCC 1709
QY 7177 TCCTTTCCCTCTGCTCAGGGCACCCAGAGCCAAAGGTTCGCGCTCCAGGACCCCTGAGTG 7236
Db 1708 TCCTTTCCCTCTGCTCAGGGCACCCAGAGCCAAAGGTTCGCGCTCCAGGACCCCTGAGTG 1649
QY 7237 CTTTCGCCACCTCTGCTGATTTACAGAGCGCATCTGCAAAATGGAGAGGCGAGTCCAC 7296
Db 1648 CTTTCGCCACCTCTGCTGATTTACAGAGCGCATCTGCAAAATGGAGAGGCGAGTCCAC 1589
QY 7297 GCCTGTGCTGACGTGGCTGGGCCACCTTTCTTGTGTCAGTCCCTAGGCGGTTTGGGG 7356
Db 1588 GCCTGTGCTGACGTGGCTGGGCCACCTTTCTTGTGTCAGTCCCTAGGCGGTTTGGGG 1529
QY 7357 ACAGGTGAGGACAGCTGCACAGAGTCTGGGCATCTACAGTGGTGACAGCAGCCACGGG 7416
Db 1528 ACAGGTGAGGACAGCTGCACAGAGTCTGGGCATCTACAGTGGTGACAGCAGCCACGGG 1469
QY 7417 CTTGTGAGACTTTTCTGGCCAGGGCAGCATCTGCCATFCCCTTCGGTGGCGATGGGA 7476
Db 1468 CTTGTGAGACTTTTCTGGCCAGGGCAGCATCTGCCATFCCCTTCGGTGGCGATGGGA 1409
QY 7477 CTGAGACCCCTCGGTGGATGGATGGCCAGAGAGGTCCTGGAGTTCCAGCCACTGG 7536
Db 1408 CTGAGACCCCTCGGTGGATGGATGGCCAGAGAGGTCCTGGAGTTCCAGCCACTGG 1349
QY 7537 CCGGCAACCTTGCTCTACCTTGCTCTCCCACTGCGCCAGTGCAGAGAGTGGCGCA 7596
Db 1348 CCGGCAACCTTGCTCTACCTTGCTCTCCCACTGCGCCAGTGCAGAGAGTGGCGCA 1289
QY 7597 TAGTGAGCCGAGAGCGGAGGCGGCGAGGCCAGAGAGCGTGGGGGAGGAGCCGGG 7656
Db 1288 TAGTGAGCCGAGAGCGGAGGCGGCGAGGCCAGAGAGCGTGGGGGAGGAGCCGGG 1229
QY 7657 AAGCCCGGCGGCGGCGCCACGCGGGAGTCCAAAGCCAGAGGAGCCCGCGCCCAAGA 7716
Db 1228 AAGCCCGGCGGCGGCGCCACGCGGGAGTCCAAAGCCAGAGGAGCCCGCGCCCAAGA 1169
QY 7717 AGCCAGACTTGGACAAGGCGCTGGGCACCGGCCAGGCTGCAGTGTGAGGACCCCGCGGA 7776
Db 1168 AGCCAGACTTGGACAAGGCGCTGGGCACCGGCCAGGCTGCAGTGTGAGGACCCCGCGGA 1109
QY 7777 AGCCTCTGGGACTGCTGCTGGGCACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGG 7836
Db 1108 AGCCTCTGGGACTGCTGCTGGGCACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGG 1049
QY 7837 TGCAGCACCCGCGCAGCATCACACCGCGGAGGCTCCAGTGTCACTTTCCAGAGTGAGA 7896
Db 1048 TGCAGCACCCGCGCAGCATCACACCGCGGAGGCTCCAGTGTCACTTTCCAGAGTGAGA 989
QY 7897 AGATGAAGGCGCATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGGCGCATCAAGC 7956
Db 988 AGATGAAGGCGCATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGGCGCATCAAGC 929
QY 7957 TGCACACTCAGGCGCAGTGCAGTGCAGTGAAGAGCAGAAAGTGTCCACCCTTAGTG 8016
Db 928 TGCACACTCAGGCGCAGTGCAGTGCAGTGAAGAGCAGAAAGTGTCCACCCTTAGTG 869
QY 8017 ACTACACTCTGCTTTCTTCCTAAGCGCGCAGCGCAAGGCGCTCTCAACTACTGGGACTTCG 8076
Db 868 ACTACACTCTGCTTTCTTCCTAAGCGCGCAGCGCAAGGCGCTCTCAACTACTGGGACTTCG 809
QY 8077 GACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 8135
Db 808 GACCGCTTGTGGGACCCAGGCTCCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 749
QY 8136 CCCCAGCCCAAGGAGGACAGGCTCACCTCTACCCAAACCTTAGTTCGCGGTCCGAGT 8195

Db 748 CCCCAGCCCAAGGGGACAGGCGCTCACCTCTACCCAAAACCTAGGTTTCCGGTCCCGAGT 689
QY 8196 ACAGTCTGTATCAAAACCCACGATTTCTCCAGCTCAGAAACCCAGGGCTCTTGCCCAAGTCG 8255
Db 688 ACAGTCTGTATCAAAACCCACGATTTCTCCAGCTCAGAAACCCAGGGCTCTTGCCCAAGTCG 629
QY 8256 TTAGAATATAGGTCCTCTTCTCCAGATCCAGCGGCCCAATGAAACCTCACGCTGGT 8315
Db 628 TTAGAATATAGGTCCTCTTCTCCAGATCCAGCGGCCAATGAAACCTCACGCTGGT 569
QY 8316 CTTAATACAGCTCTTTAAAGGCCAGCCCTAGAAACCCAAAGCTCTCTCTCGGAACCGC 8375
Db 568 CTTAATACAGCTCTTTAAAGGCCAGCCCTAGAAACCCAAAGCTCTCTCTCGGAACCGC 509
QY 8376 TCACCTAGAGCCAGACCAACGTTACTCAGGCTCTCCAGGCTCTCTCCAGCTCTGAGGCTTT 8435
Db 508 TCACCTAGAGCCAGACCAACGTTACTCAGGCTCTCTCCAGCTCTCTCCAGCTCTGAGGCTTT 449
QY 8436 CACCTTTAACCCAAAGGAGCAGAGTCCACCTCCAGCGCGGGGAGCTTAGGACCACTCA 8495
Db 448 CACCTTTAACCCAA-GSAGCACAGGTCCCACCTCCAGGCC-GGAGGCTTAGGACCACTCA 391
QY 8496 GCCCTAGGAGTATATTTCCGCACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTG 8555
Db 390 GCCCTAGGAGTATATTTCCGCACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTG 331
QY 8556 CCCCATAACTCTCAGTCTGCTTCCAGAAATTTGGAATCTAGTTTCTCTCTCTCGTA 8615
Db 330 CCCCATAACTCTCAGTCTGCTTCCAGAAATTTGGAATCTAGTTTCTCTCTCTCGTA 272
QY 8616 TCCGAGTCTGGGACACAAACCTCCGCCCCAGCCTATGAGCATCTGAGCCCCGCCCTC 8675
Db 271 TCCGAGTCTGGGACACAAACCTCCGCCCCAGCCTATGAGCATCTGAGCCCCGCCCTC 212
QY 8676 TTCTGACGAACTGGCCCCGGATCAGAGCAGGACCTCCCTTCGACCTCTGGGAACCT 8735
Db 211 TTCTGACGAACTGGCCCCGGATCAGAGCAGGACCTCCCTTCGACCTCTGGGAACCT 152
QY 8736 CCGAGAGTCCAGCCCATCTCGGAGCATCCGGAGGAAATCTGCAGAGGGGTTAGGAGTG 8795
Db 151 CCGAGAGTCCAGCCCATCTCGGAGCATCCGGAGGAAATCTGCAGA-GGGTTAGGAGTG 93
QY 8796 GGTGACAAGAGCCTGATCT 8855
Db 92 GGTGACAAGAGCCTGATCT 33
QY 8856 AGATGATACATTTTGTAAAAAATAATAA 8887
Db 32 AGATGATACATTTTGTAAAAAATAATAA 1

RESULT 6
PCT-US01-14827-388
; Sequence 388, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 388
; LENGTH: 2770
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2604)..(1804)
; OTHER INFORMATION: 100% homologous to Homo sapiens C33 antigen-type III integral
; OTHER INFORMATION: membrane protein,accession number S48196.Smith-Waterman Score=

OTHER INFORMATION: 1413.
PCT-US01-14827-388

Query Match 15.58; Score 1425.2; DB 1; Length 2770;
Best Local Similarity 99.1%; Pred. No. 7.6e-22;
Matches 1433; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
QY 2288 ATGGGGCTGAAGCGCGCCAGAACAGCGTGTTCCTGGCTGCCTCCATCGACGAGCTGGTG 2347
Db 1 atg999ctgaaggccgcccagaagacgctgtctcccgctgscgtccatcagcagctgggtg 60
QY 2348 CGCCTGTTTGTGCGAGCTGGCGCGAGAGAGCGGACCTGCTGCTCTTTCCTTTGCTG 2407
Db 61 cgcctgtttgtcgcagagtcggccgagagagcgcggaacctggctccttctccttggtg 120
QY 2408 CTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCGCTATCCCTACCAAGCTTCCGAG 2467
Db 121 ctg99cttcgtggagcattttctgctgtcaaccgctcatcctccatcccaacggtcccgag 180
QY 2468 CTCACCTTCCAGCCAGCGCCCGCCCGACCGCGCTGGCGGCTTCACTACTTTCCTGGTG 2527
Db 181 ctacacctccagccagcccgcccgcccgacccgctcgtg9ccctcaacctacttctcccgctg 240
QY 2528 GCCGACCTGTCTATCATGCGCGCCCTCTATGCCGCTTACCGCCCGACGATCCGAGCGCC 2587
Db 241 gccgacctgtctatcatcgcgcgcctctctatgcccgccttcaaccgcccagatcccgagcgcc 300
QY 2588 GTCGACCTGTCCTTATCTTCCTCGAGAGGGGGTGTCTCCAGCCGCTGAGCTGGTGAAGAAG 2647
Db 301 gtcgacctgtccctctatctctcgagaagg99gtctccagccgtgagctg9gtgaagaag 360
QY 2648 GTCTCCGATGTATATGAACAGCGCTCAGCGCTCTACTTCAAGGATCGGGCGCCACATC 2707
Db 361 gtctcgatgtcatatgaaacagcctcagcgcctcctacttcaaggatcgggcccacatc 420
QY 2708 CAGTCCCTCTTACGCTTCATCACAGGTTGGAGCCAGTAGTGGGAATCTTATCCATGAC 2767
Db 421 cagtccctcttcagcttcacaggttgagcccgtagtg9ggaatcttatccatgac 480
QY 2768 CCAGTCTTCAAAACCTCCATGTTTACAGAACCTTTTAAAGACTGTAAGCTTGTGA 2827
Db 481 ccaactctcctcaaacctccatg9ttacagaaaccttttaagaaactgtaagccttgga 540
QY 2828 GGTTCGGCAGGTGTATTTCTCTTCTTTCAGTTGGGAACCTGAAGCCAGAGGGGAAA 2887
Db 541 ggttcgcgagtg9ttatcttctcttctgagttgggaactgaagccagagaggggaaa 600
QY 2888 TGATATGCCAAAGTCACACAGCGGATGCGAGGGCTGGAGTGAAGCCTGATCTGGCT 2947
Db 601 tgatatgccaagtcaacacggtgatggcaggtggaagtgaagcctgatcacttggt 660
QY 2948 CCAATATCAACCTCACTCTGCGCCCTCAGACCCCGCCCTTGCACCTGAACAGCTA 3007
Db 661 ccaaatcataacctcaacctctgcccctctcagccccccaccttgccactgaaacagcta 720
QY 3008 CAGGAGTTCTAAGCATGAGACACAGAGCGCGCAGAGATTAGGGGCAAGAAATGAA 3067
Db 721 caggagttctaagcatgagacacagagggcgagcagatgtaggggggcaagaagatgaa 780
QY 3068 ATGGGCTGCATTTGAGCGCAGTTAAACAAATAATGGCTATGAAGATTTTTTTTTT 3127
Db 781 att999ctgcatgttgggcaggttaaacaaaataatggtatgaagattttttttttt 840
QY 3128 TTTTTCGAGACAGGGTCTCACTCTGTCCCGCAGGCTGGAGTGGCTGTGATCATGG 3187
Db 841 tttttttgagacaggggtctcactctgtcccccagggctgagtg9gtgtgatcatg9 900
QY 3188 CTCAGTGCAGCGCTCAGTCTCCCTGGGCTCAGAGATTCCTCCAAACCTCAGCCTCCTGAGTAG 3247
Db 901 ctcactgcagcctcagctcctcctg9gtcagagatccctccaaacctcagcctcctgagtag 960
QY 3248 CTGAGAGTACAGGCATGACCGTGGTGTGCTTAATTTTTTTTGTATTTTTTTGTAGAGAT 3307
```

```
Db 961 ctgagagtcagggcatgcaccgtggtgctggttaatttttttttttttttttttttttttttt 1020
QY 3308 GGTCTCTCACTATGTGCCCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATCTGCCCGC 3367
Db 1021 ggtctcactatgtgcccagactggtcttgaactcttggctcaagtgtatcttgcocgc 1080
QY 3368 CTCAGTCTCCCAAATGCTGGGATTTACAGGTGTAGCCACCGCAACTGGTGGCCTTATGAAA 3427
Db 1081 cteagtctccaaatgctgggtattacagggtgtagccaccgcaactggtg9ccctatgaaa 1140
QY 3428 ATTTTTCCTTTTTCAGACGGCTCTCACTCTCTGCGCCAGGCTGGAGTGCAGTGTGTC 3487
Db 1141 atttttttttttttcagacggctctcaactctgccccag9ctg9agtgacgtg9tgc 1200
QY 3488 AATCTCGGCTCACTGCAAGCTCTGCCTCCTCTTTCATGCCATTTCTCTGCTCCTGCTT 3547
Db 1201 aatctcgctcaactgcaagctctgctcctctcttctatgcatctcctcctcctcctgct 1260
QY 3548 CAGCTCTCTAGTAGTGGGACTACAGGAGCTCCACCATGCGCTGGCTAAATTTTTTTTTT 3607
Db 1261 cagctcctgagtagctg9ggaactacaggtgccccgaccgccccg9ctaaatttttttt 1320
QY 3608 GGATTTTTCAGTAGACGAGGTTTTCACCATGTTAGCCAGGATGGTCTCGATCTCTTGACC 3667
Db 1321 gtatttttagtagagaggggtttccacctgttagccaggtggtctcgatctcctcagcc 1380
*QY 3668 TCGNGATCCGCGCCCTTGGCTTCCCAAAGTCTGGGATTTACAGCGTGAAGCCGAC 3727
Db 1381 tctgtagtccaccgctcctg9ccctcccaaa9tctg9g9attacag9ctg9g9ccccc9gc 1440
QY 3728 CTGGTC 3733
Db 1441 ccg9cc 1446
```

RESULT 7

PCT-US01-14827-5297/c
; Sequence 5297, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 5297
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2121)...(1651)
; OTHER INFORMATION: 98% homologous to Homo sapiens menin, accession number
; OTHER INFORMATION: U93237, Smith-Waterman Score=783.
PCT-US01-14827-5297

Query Match 14.9%; Score 1365.2; DB 1; Length 2121;
Best Local Similarity 98.1%; Pred. No. 1.3e-20;
Matches 1413; Conservative 0; Mismatches 23; Indels 4; Gaps 3;

```
QY 2288 ATGGGGCTGAAGCGCGCCAGAACAGCGTGTTCCTGGCTGCCTCCATCGACGAGCTGGTG 2347
Db 2121 ATGGGGCTGAAGCGCGCCAGAACAGCGTGTTCCTGGCTGCCTCCATCGACGAGCTGGTG 2062
QY 2348 CGCCTGTTTGTGCTGGCGCGAGAGGAGCGGACCTGGTGTCTCTTTTCTCTTTGGTG 2407
Db 2061 CGCCTGTTTGTGCTGGCGCGAGAGGAGCGGACCTGGTGTCTCTTTTCTCTTTGGTG 2002
```

Qy	2408	CTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCGTCACTCCATCCACAGTFTCCCGAG	2467
Db	2001	CTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCGTCACTCCATCCACAGTFTCCCGAG	1942
Qy	2468	CTCACCTTCCAGCCAGCCGCCGCCGCGCTGGCGGCTCACTACTTTCCTCCGGTG	2527
Db	1941	CTCACCTTCCAGCCAGCCGCCGCCGCGCTGGCGGCTCACTACTTTCCTCCGGTG	1882
Qy	2528	GCGGACCTGTCTATCATCGCCGCCCTCTATGCGCCGCTTCACCGCCAGATCCAGAGCGCC	2587
Db	1881	GCGGACCTGTCTATCATCGCCGCCCTCTATGCGCCGCTTCACCGCCAGATCCAGAGCGCC	1822
Qy	2588	GTGAGCTGTCCCTCTATCTCTGAGAGAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAG	2647
Db	1821	GTGAGCTGTCCCTCTATCTCTGAGAGAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAG	1762
Qy	2648	GTCTCGAGTGTATATGGAACAGCTCTAGCGCGCTCTACTTCAAGGATCGGGGCCACATC	2707
Db	1761	GTCTCGAGTGTATATGGAACAGCTCTAGCGCGCTCTACTTCAAGGATCGGGGCCACATC	1702
Qy	2708	CAGTCCCTCTTCAGCTTTCATCAGAGTGTGGAGCCAGTAGTGGGAATCTTATCCATGAC	2767
Db	1701	CAGTCCCTCTTCAGCTTTCATCAGAGTGTGGAGCCAGTAGTGGGAATCTTATCCATGAC	1642
Qy	2768	CCACTTCTTCAAAACCCCTCCATGGTTTACAGAACCCCTTTTAAAGAACTGTAAAGCCTTGTGA	2827
Db	1641	CCACTTCTTCAAAACCCCTCCATGGTTTACAGAACCCCTTTTAAAGAACTGTAAAGCCTTGTGA	1582
Qy	2828	GGTTGGCAGGTGTATATTTTCCTCTTTGCAGTTGGGAACCTGAAGCCAGAGGGGAAA	2887
Db	1581	GGTTGGCAGGTGTATATTTTCCTCTTTGCAGTTGGGAACCTGAAGCCAGAGGGGAAA	1522
Qy	2888	TGATATGCCAAAGTCAACACAGGCCTGGCAGGGCTTGAAGTGAAGCCGTGATCACTTGGCT	2947
Db	1521	TGATATGCCAAAGTCAACACAGGCCTGGCAGGGCTTGAAGTGAAGCCGTGATCACTTGGCT	1462
Qy	2948	CCAAATCATCAACCTCACTCTGCCCCCTCAGACACCCCAACCTTGGCAGCTGAACAGCTA	3007
Db	1461	CCAAATCATCAACCTCACTCTGCCCCCTCAGACACCCCAACCTTGGCAGCTGAACAGCTA	1402
Qy	3008	CAGGAGTTCTAAGCATGAGACACAGAGGGCGCAGCAGATTTAGGGGCGAAGAGATGAA	3067
Db	1401	CAGGAGTTCTAAGCATGAGACACAGAGGGCGCAGCAGATTTAGGGGCGAAGAGATGAA	1342
Qy	3068	ATTGGGCTGATTTGAGCAGTTAAACAAAATAATGGCTATGAAGATTTTTTTTTTTTT	3127
Db	1341	ATTGGGCTGATTTGAGCAGTTAAACAAAATAATGGCTATGAAGATTTTTTTTTTTTT	1282
Qy	3128	TTTTTTTGGACAGGCTCTCACTCTGTCCCCAGGCTGGAGTGGAGTGGTGTGATCATGG	3187
Db	1281	TTTTTTTGGACAGGCTCTCACTCTGTCCCCAGGCTGGAGTGGAGTGGTGTGATCATGG	1222
Qy	3188	CTCAGTGCAGCCTCAGCTCTCCCTGGGCTCAGAGATCTCTCAACCTCAGCCTCTGAGTAG	3247
Db	1221	CTCAGTGCAGCCTCAGCTCTCCCTGGGCTCAGAGATCTCTCAACCTCAGCCTCTGAGTAG	1162
Qy	3248	CTGAGAGTACAGGCACTGCACCGTGTG - CTGGTTAAATTTTTTGTATTTTTTCTACAGA	3306
Db	1161	CTGAGAGTACAGGCACTGCACCGTGTG - CTGGTTAAATTTTTTGTATTTTTTCTACAGA	1102
Qy	3307	TGGTGTCTCACTATGTGGCCCAAGCTGGTCTTGAACCTCTGGGCTCAAGTGATCTGCCG	3366
Db	1101	TGGTGTCTCACTATGTGGCCCAAGCTGGTCTTGAACCTCTGGGCTCAAGTGATCTGCCG	1042
Qy	3367	CCTCAGTCTCCAAATGCTGGGATTCAGAGTGTGAGCCACCGCAACTGGTGGCCCTATGAA	3426
Db	1041	CCTCAGTCTCCAAATGCTGGGATTCAGAGTGTGAGCCACCGCAACTGGTGGCCCTATGAA	982
Qy	3427	AAATTTTTTTTTTTTTTTCAGACCGGCTCTCACTCTGTGCCAGGCTGGAGTGCAGTGGT	3486
Db	981	AAATTTTTTTTTTTTTTTCAGACCGGCTCTCACTCTGTGCCAGGCTGGAGTGCAGTGGT	922
Qy	3487	CAATCTCGGCTCACTGCAAGCTCTGCCTCTCTTTCATGCACTCTCCTGCCCTCTGCC	3546

Db	921	CAATCTCGGCTCACTGCAAGCTCTGCTCTCTCTCTTTCATGCCATTTCTCTGCCTCTGCTGCC	862
Qy	3547	TCAGCCTCTGAGTAGTGGGACTACAGAGGCTGCCACCATGCTGCTAAATTTTTTTTT	3606
Db	861	TCAGCCTCTGAGTAGTGGGACTACAGAGGCTGCCACCATGCTGCTAAATTTTTTTTT	802
Qy	3607	TGATTTTTAGTAGAGAC-GAGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCTCGA	3665
Db	801	TGATTTTTAGTAGAGACAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAATCTCTGA	742
Qy	3666	CCTC-CTGATCCGCCCGCTTGGCTCCCAAGTCTGGGATTAACAGCCTGAGCCACC	3723
Db	741	CCTCAGGTGATCCGCCCGCTTGGCTCCCAAGTCTGGGATTAACAGCATGAGTACGC	682
RESULT 8			
US-09-948-941-5			
; Sequence 5, Application US/09948941			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL000788			
; CURRENT APPLICATION NUMBER: US/09/948,941			
; CURRENT FILING DATE: 2001-09-10			
; PRIOR APPLICATION NUMBER: 60/231,328			
; PRIOR FILING DATE:			
; NUMBER OF SEQ ID NOS: 12618			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 2772			
; TYPE: DNA			
; ORGANISM: Human			
US-09-948-941-5			
Query Match 14.3%; Score 1312.2; DB 5; Length 2772;			
Best Local Similarity 96.5%; Pred. No. 1.1e-19;			
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps			
Qy	7499	GGATGGCCAGACGAGGTCTGTGAGTTCCAGCCACTGGCGCGCAACCTTGCTCTCACCTT	7558
Db	1382	ggagggcagtcctcaagcctgtctgcactgtggcttggccactcttctgtgcagtcctt	14411
Qy	7559	GCTCTCCCCACTGCGCCAGGTGCGGCAGAAAGTGCGCATAGTGAGCCGAGAGCGCCGAGGC	7618
Db	1442	agggcgtttgaggagcaggtgcgcagaaggtgcgcatagtgagccgagagcgccgagc	1501
Qy	7619	GGCGAGCGCCGAGAGCGGTGGGGCCGAGAAAGCCCGGGAAGCCCGCGCGGGCGGCCACG	7678
Db	1502	ggcgcagcgcagagccgttggggcaggaagcccggaaggccgcgcggcgccgcacg	1561
Qy	7679	GCGGGAGTCCAAGCCAGAGAGGCCCGCGGCCCAAGAACCCAGCACCTGGACAAGGGGCT	7738
Db	1562	gcgggagttccaaagcagagagagcccccgcgcgcgaagcagcacttggaacagggcct	1621
Qy	7739	GGCAGCCGCCAGGGTGCAGTGTGTCAGGACCCCCCGGAAGCCTCTCGGACATGTCTGCTGG	7798
Db	1622	gggcacccgcaggggtgcaggttcaggagcccccccggaagcctcttgggactgtctgtg	1681
Qy	7799	CACAGCCGAGGCGCTGAAGTGGCAGCAGGCTCAGGTGCCAGCACCAGCCGAGCATACC	7858
Db	1682	cacagcccgagccctgaaggttggcagcagcgtctcaggltccacagcccccgagcatcacc	1741
Qy	7859	ACCCCGGAGGGTCCAGTGTCTACTTTCCAGAGTCAGAAAGATGAAGGGCATGAAGGAGCT	7918
Db	1742	accgcgaggggtccagtgctcaactttccagtagaagaatgaaaggcatgaagagact	1801
Qy	7919	GCTGGTGCCACCAAGATCAACTCTGAGCGGCATCAAGCTGCAACTCACGGCACATCGCA	7978
Db	1802	gctggtgcccacaaagatcaacttcagcgcctcaagcttgcaacttcagcgcagcagtcgca	1861

Qy 7979 AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACTCTGTCTTTCTCAA 8038
Db 1862 agtcagatgaagaagcagaaagtgtccacccctagtgaactacactctgtcttctctcaa 1921
Qy 8039 GCGCAGCGCAAAAGCCTCTCAACTACTTGGGACCTTGGGACCCAGGC 8098
Db 1922 gcgagcagcaaaagcctctgaactactggagactctggagccgcttctggagcccgagc 1981
Qy 8099 TCCGCCCTTAGTCCGCCCACTCTGAGCCCATGTTCTGCCGCCAGCCCAAGGGGACAGGCC 8158
Db 1982 tccgcttagtcccccaactctgagccatgttctgtccccagcccaaaagggagacagcc 2041
Qy 8159 TCACCTCTACCAAAACCCCTAGTCTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
Db 2042 tcacctctacccaaacccctaggttctccggtcccgagtacagctgtgtatcaaaacccagat 2101
Qy 8219 TTTCTCCAGCTCAGAACCCAGGCTCTGCCCCAGTCTGTAGATATAGGCTCTTCTCCC 8278
Db 2102 ttctccagctcagaacccagggctctgccccagctgttagaaatataggtctctctccc 2161
Qy 8279 AGAATCCAGCGCGGCAATGGAACCTCACGCTGGTCTCAATTACCAGTCTTTAAAGGC 8338
Db 2162 agaattccagcgccgcaatggaaacctcaagctgggtcttaattaccagctctttaaaggc 2221
Qy 8339 CCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 8398
Db 2222 ccagcccttagaaaccccaagctctctcctcggaacgctcacttagagccagaccagctt 2281
Qy 8399 ACTCAGGCTCTCCAGCTGTAGGAGCTGAGGCTTTACCCCTTAACCCAGGAGGACACA 8458
Db 2282 actcagggctctccagctctgttaggagctgaggttccaccttaaccccaaggagagcaca 2341
Qy 8459 GGTCCCACTCTCAGCCCGGAGGAGCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCA 8518
Db 2342 ggtcccaactccagcccgaggagcctaggaccactcagcccttaggagtatatttcgca 2401
Qy 8519 CTTCAGAAATCCATATCTTGGAAATCCAAGCTCCCTGCCCAATAACTTCAGTCTGCT 8578
Db 2402 cttcagaattccatctcttgcgaatccaaagctcctcgcccccaataaacttcagctcctgct 2461
Qy 8579 TCCAGAAATTTGAAATCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
Db 2462 tccagaatttggaaatcctctagtttctctctctctctctctctctctctctctctctctct 2521
Qy 8639 CCGCCCCCAGCCTATGAGCAATCTGAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8698
Db 2522 ccgccccagcctatgagcactctgagcccgccctctctctctgacgaaactggcccgga 2581
Qy 8699 TCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGG 8758
Db 2582 tcagagcaggacctctctctcgaacctctggaaacctcccaagaggtccagcccatctcgg 2641
Qy 8759 AGCATCCCGAGGAAATCTGAGAGGGGTTAGGAGTGGGTGACAGAGCCTGATCTCTTC 8818
Db 2642 agcatcccgaggaaatctgcagagggggttaggagtgagggtgagcagagcctgactctctc 2701
Qy 8819 CTGTTTGTACATAGATTTATTTTCTAGTCTCCAAAGAAAGATGAATACATTTTGTAAAAA 8878
Db 2702 ctgtttgtacatagattttattttctagttcccaagaaagatgaatacatattttgttaaaa 2761
Qy 8879 AAATATANA 8887
Db 2762 aaaaaaaa 2770

RESULT 9

US-09-380-337-1

; Sequence 1, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Setara C.
; Guru, Siradanahalli C.
; Manickam, Pachiaipan
; Collins, Francis S.

Emmert-Buck, Michael R.
Debelenko, Larisa V.
Lubensky, Irina A.
Liotta, Lance A.
Agarwal, Sunita K.
Spiegel, Allen M.
TITLE OF INVENTION: MEN1, the Gene Associated With Multiple Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses Thereof
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/380,337
FILING DATE: 09-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/040,269
FILING DATE: 05-Mar-1997
APPLICATION NUMBER: WO PCT/US98/04258
FILING DATE: 04-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: Lockyer, Jean M.
REGISTRATION NUMBER: 44,879
REFERENCE/DOCKET NUMBER: 015280-315100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2772 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 111..1940
OTHER INFORMATION: /product= "human menin"
FEATURE:
NAME/KEY: exon
LOCATION: 1..87
FEATURE:
NAME/KEY: exon
LOCATION: 88..555
FEATURE:
NAME/KEY: exon
LOCATION: 556..764
FEATURE:
NAME/KEY: exon
LOCATION: 765..893
FEATURE:
NAME/KEY: exon
LOCATION: 894..934
FEATURE:
NAME/KEY: exon
LOCATION: 935..1022
FEATURE:
NAME/KEY: exon
LOCATION: 1023..1159
FEATURE:
NAME/KEY: exon
LOCATION: 1160..1295

```

;
; NAME/KEY: exon
; LOCATION: 1296..1460
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 1461..2764
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-337-1

Query Match      14.3%; Score 1312.2; DB 5; Length 2772;
Best Local Similarity 96.5%; Pred. No. 1.1e-19;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7499 GGATGGCCAGAGCAGGGTCTGTGAGTTCCAGCCACTGCGCGGCAACCTTCTCTCACCTT 7558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1382 GGAGGGCAGTCCCACGCGCTGTGTGACGTGGCGTGGCCACCTTCTTGTGTCAGTCCCT 1441

QY 7559 GCTCTCCCACTGGCCAGGTGGCGGAGAGGTGCGCATAGTAGCGGAGAGCCGAGGC 7618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 AGCCCGTTTTGAGGGACAGGTGGCGGAGAGGTGCGCATAGTAGCGGAGAGCCGAGGC 1501

QY 7619 GCGGAGGCGGAGAGCGGTGGGGCGGAGGAGCCCGGAGGCGCGGGGGCCACG 7678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1502 GCGCGAGGCGGAGAGCGGTGGGGCGGAGGAGCCCGGAGGCGCGGGGGCCACG 1561

QY 7679 GCGGGAGTCCAAAGCAGAGAGCGCCCGCGCCCAAGAGCCAGCACTGGACAGGGCCCT 7738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1562 GCGGGAGTCCAAAGCAGAGAGCGCCCGCGCCCAAGAGCCAGCACTGGACAGGGCCCT 1621

QY 7739 GGCACCGGCGAGGGTGCAAGTGTGAGAGCCCGCCCGGAGCCCTCTGGGACTGTGCTGG 7798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1622 GGCACCGGCGAGGGTGCAAGTGTGAGAGCCCGCCCGGAGCCCTCTGGGACTGTGCTGG 1681

QY 7799 CACAGCCCGAGGCGCTGAAGTGGGAGCAGCGCTCAGGTGCCAGCACCAGCCAGCATACC 7858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1682 CACAGCCCGAGGCGCTGAAGTGGGAGCAGCGCTCAGGTGCCAGCACCAGCCAGCATACC 1741

QY 7859 ACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAGAGATGAAGGATGAAGGAGCT 7918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1742 ACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAGAGATGAAGGATGAAGGAGCT 1801

QY 7919 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 7978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1802 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 1861

QY 7979 AGTGCAGATGAAGAACAGACAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCCTCA 8038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1862 AGTGCAGATGAAGAACAGACAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCCTCA 1921

QY 8039 GCGGAGCGCAAGGCGCTCTGAAGTACTTGGGACTTTCGGACCGCTTGTGGGACCCAGGC 8098
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1922 GCGGAGCGCAAGGCGCTCTGAAGTACTTGGGACTTTCGGACCGCTTGTGGGACCCAGGC 1981

QY 8099 TCGGCGCTTAGTCCCGCAACTCTGAGCGCCATGTGTTCTGCCCGCCAGCCAAAGGGACAGGC 8158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1982 TCGGCGCTTAGTCCCGCAACTCTGAGCGCCATGTGTTCTGCCCGCCAGCCAAAGGGACAGGC 2041

QY 8159 TCACCTCTACCCAAAGCCCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2042 TCACCTCTACCCAAAGCCCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 2101

QY 8219 TTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCGAGTGTGAGATATAGGTCTCTCTCCC 8278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 TTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCGAGTGTGAGATATAGGTCTCTCTCCC 2161

QY 8279 AGAATCCCGAGCGGCAATGGAAACCTCAGCGTGGGTCTTAATACCAAGCTTTTAAAGGC 8338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2162 AGAATCCCGAGCGGCAATGGAAACCTCAGCGTGGGTCTTAATACCAAGCTTTTAAAGGC 2221

QY 8339 CCAGCGCCCTAGAAACCAAGCTCTCTCGGAACCCGCTACCTAGAGCCAGACCAAGCTT 8398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2222 CCAGCGCCCTAGAAACCAAGCTCTCTCGGAACCCGCTACCTAGAGCCAGACCAAGCTT 2281
```

```

QY 8399 ACTCAGGGCTCCTCCAGCTTGTAGGAGCTTGAGGTTTCAACCCTTAACCCAAAGGAGCACA 8458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2282 ACTCAGGGCTCCTCCAGCTTGTAGGAGCTTGAGGTTTCAACCCTTAACCCAAAGGAGCACA 2341

QY 8459 GTTCCCACTCCAGCGCGGGAGCCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCGCA 8518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2342 GTTCCCACTCCAGCGCGGGAGCCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCGCA 2401

QY 8519 CTTCAGAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAATATACTTCAGTCTCTGCT 8578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2402 CTTCAGAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAATATACTTCAGTCTCTGCT 2461

QY 8579 TTCAGAATTTGGAATCTCTAGTTTCTCTCTCTCGTATCCCGAGCTCTGGGACACAAACT 8638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2462 TTCAGAATTTGGAATCTCTAGTTTCTCTCTCTCGTATCCCGAGCTCTGGGACACAAACT 2521

QY 8639 CCGCCCCAGCCTATGAGCATCTGAGCCCCCGCCCTTCTCTGACGAAACTTGCCCCCGGA 8698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2522 CCGCCCCAGCCTATGAGCATCTGAGCCCCCGCCCTTCTCTGACGAAACTTGCCCCCGGA 2581

QY 8699 TCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCACGCCATCTCGG 8758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2582 TCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCACGCCATCTCGG 2641

QY 8759 AGCATCCCGGAGGAAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCTGTATCTCTTC 8818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2642 AGCATCCCGGAGGAAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCTGTATCTCTTC 2701

QY 8819 CTGTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGAATACATTTTGTAAAAA 8878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2702 CTGTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGAATACATTTTGTAAAAA 2761

QY 8879 AAATATAAA 8887
    ||| ||| |||
Db 2762 AAAAAAAA 2770
    ||| ||| |||

RESULT 10
US-09-948-941-125
; Sequence 125, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000798
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 2767
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-125

Query Match      13.7%; Score 1256.4; DB 5; Length 2767;
Best Local Similarity 96.3%; Pred. No. 1.4e-18;
Matches 1339; Conservative 0; Mismatches 46; Indels 5; Gaps 5;

QY 7499 GGATGGCCAGAGCAGGGTCTGTGAGTTCCAGCCACTGGCGGCAACCTTCTCTCACCTT 7558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1382 ggaaggcagtcacccacgctgtgtgcactggtggcgacaccccttcttgcagtcct 1441

QY 7559 GCTCTCCCCACTGCGCCAGGTGCGGACAGAGTTCGCGATAGTGCAGCGGAGAGCGGAGGC 7618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 aggccttttgaggagacagtgcggaagaggtgcacatagtgagccgagagccgagagc 1501

QY 7619 GCGCGAGCGCGGAGGCGGTGGGGCGGAGGAAAGCCCGGAGGCGCGGGGCCACG 7678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


[illegible]

RESULT 12

US-60-278-561-1743

US-00-278-301-1743
; Sequence 1743, Application US/60278561

; GENERAL INFORMATION:

; APPLICANT: MORRIS, MacDONALD

APPLICANT: Lal, Preeti

; APPLICANT: Diep, Dinh

```

; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
;
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and SingleNucleotide

```

```

; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 1743
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; US-60-278-561-1743

```

Query Match	13.4%;	Score 1231.4;	DB 8;	Length 3534;	1
Best Local Similarity	96.2%;	Pred. No. 3.4e-18;			
Matches 1325;	Conservative	0;	Mismatches 46;	Indels	6;
Gap's					
Qy	7499	GGATGGCCAGACGACGGTCTCTGGAGTTCCAGCCACTGGCCGGCAACCTTGCTCTCACCTT	7558		
Db	2162	ggaggcaggtccacgcctgtgctgcacgtgggtctgggccacctttctt g gaatccct	2221		
Qy	7559	GCTCTCCCCACTGGCCCCAGGTGGCGCAGAAAGTGTGCGCATAGTGTAGCCGAGAGCCGAGGC	7618		
Db	2222	aggcgtttttagggacaggtgcggcagaaggtgcgcatagt gagccgagagcgcgaggc	2281		
Qy	7619	GGCGAGGCGCAGAGCCGTGGGGCGAGAAACC CGGAAGCCGGCGCGGGCGGCCACG	7678		
Db	2282	ggccgagcccgagagccgttggggcgagagcccggaagccggcgggggccccacg	2341		
Qy	7679	GCGGGAGTCCAAGCCAGAGAGACCCCGCGGCCCAAGAACCCAGCACATGTGAACAGGCGCT	7738		
Db	2342	gcgggagttcaagccagagagagccccgcgcgccaaagccagcact ggacaagggcct	2401		
Qy	7739	GGGCACGGCCAGGGTGCTAGTGTACAGGA-CCCCCGCGGAAGCCTCTCTGGGACTGTGCGTG	7797		
Db	2402	gggcacggccaggg gcaggt tcagagaccgcccggaagcctcctgggact gcctg	2461		
Qy	7798	GCACAGCCGACGAGCCCTGAAGGTGGCAGCACGGCTCAGTGTGCAGCACCCGACGATCAC	7857		
Db	2462	gcacagcccgagccctggaaggttggcagcacaggtctcaggtgccagcaccgcgagca tcac	2521		
Qy	7858	CACGCGCGAGGTCCAGTGCTCTACTTTCCAGAGTGAGAAGATGAAGGCATGAAGGAGC	7917		
Db	2522	cacgcgcggaggttccag tgcactttccagag tgagaagatgaaaggc atgaaggagc	2581		
Qy	7918	TGCTGGTGGCCACCAAGATCAACTTCAGCGCCATCAAGCTGCAACTCAGGCACAGTCGC	7977		
Db	2582	tgc tgg tggccaccagatacaactgagcgcatacaagctgcaactcagcgcag tcgc	2641		
Qy	7978	AAGTGCAGATGAAGAAGCAAGTGTCCACCCCTAGTGACTTACACTCTGTCTTTTCTCA	8037		
Db	2642	aag tgagat tgaaagacgaaagtgtccacccttag gactaacact gtctt tctca	2701		
Qy	8038	AGCGGCAGCCAAAGGCGCTGTGAACTACTTGGGGACTTGGACCCGCTTGTGGGGACCCAGG	8097		
Db	2702	agcggcagcgcaaggcctctgaactacttggggaactcggacgcgt tgggggccccag	2761		
Qy	8098	CTCGG-CCTTTAGTCCCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGG	8156		
Db	2762	ctccgcctt tag tcccccaactctgagccca t tctgccccccagccccaaagg gacag	2821		
Qy	8157	CCTCACCTCTACCCAAACCOCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACG	8216		
Db	2822	cctaactctacccaaaccttagttcccg tcccg tcccgag t cag tct g t c a c a c c c a c g	2881		
Qy	8217	ATTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTGGTTAGATATAGTCTCTTCTC	8276		
Db	2882	atttctccagctcagaaccagggcctctgccccagtcgt tagaata tag t c t c t c t c c	2941		
Qy	8277	CCAGAAATCCCAAGCGGGCAATGAAACCTCAGCGTGGGTCTTAATTACAGTCTTTAAAG	8336		

2942	Db		ccgagaatcccgccgagcccaatggaaacctcagcgcgggtccctaaatccacagtccttaag	3001
8337	Qy		GCCAGCCCTAGAAACCAAGCTCCTCTCGGAACCGCTCACCTAGAGCCAGACCAACG	8396
3002	Db		gccagccctagaaccacaagctcctcctcggaccgctcacctagagccagaccaacg	3061
8397	Qy		TTACTCAGGGCTCCTCCCAAGCTGTGTAGAGCTTGAGGTTTACCCTTAACCCAGGGAGCA	8456
3062	Db		ttactcagggtcctccagctgttaggagctgaggtttcaccccttaaccaca-ggagca	3120
8457	Qy		CAGGTCCCACTCCACGCCGGGAGGCTAGGACCACTCAGGCCCTFAGGAGTATATTTCCG	8516
3121	Db		caggtcccaacctcagccc-gggagccttaggaccaactcagccccctaggagtatattccg	3179
8517	Qy		CACTTCAGAATTCCTATCTTCGGAATCCAAGCTCCCTGCCCCAAATTAACCTCAGTCTG	8576
3180	Db		cacttcagaattcca tatcttgogaatccaagctccctgccccaaataaacttcagtcctg	3239
8577	Qy		CTTCCAGAATTGGAAATCCTAGTTTCTCTCTCTTCTGATCCCGAGTCTGGGACACAAAA	8636
3240	Db		c-ccagaatttggaaatcctagtttccctctctctcg tatccccagctcgggacacaaaa	3298
8637	Qy		CTCGGCCCCAGAGCTATGAGCATCTGTAGCCCGCCGCTCTTCTCTGACCAAACTGCCGCCG	8696
3299	Db		ctcgccccagcctatgagcactcagccccgcctctcttcgacgaaactggccccg	3358
8697	Qy		GATCAGAGCAGGACCTCCCTTCGGACCTCTCGGAACTCTCCAGAGGTCAGGCCATCTC	8756
3359	Db		gatcagcagggacctcccttcgacccctctgggaacctccagaggtccagggcccatctc	3418
8757	Qy		GGAGCATCCGGAGGAAATCTCAGAGGGGTTAGAGTGGGTGACACAGGCTGATCTCT	8816
3419	Db		ggagcatcccggaggaaa tctcaga -gggttaggagtgggtagcaagagcctgactctc	3477
8817	Qy		TCCTGTTTGTACATAGATTTATTTTTCAGTTCCTCCAGAAAGATGAATACATTTTGT	8873
3478	Db		tctgtttgtacatagaattatctttccagttccagagaaataaatacatttgtt	3534

```

RESULT 13
US-60-278-561-1743/c
; Sequence 1743, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Deep, Diah
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 1743
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
US-60-278-561-1743

```

	Query Match	7.6%	Score 698.6;	DB 8;	Length 3534;
	Best Local Similarity	95.5%;	Pred. No. 6.4e-08;		
	Matches 719;	Conservative 0;	Mismatches 34;	Indels 0;	Gaps 0;
Oy	2229	CCTGCGCCAGACCTCCTCCTCCCGGGTTGCCCTTGCAAGCGCGCGGCCACCGCGCCGCGCGCA	2288		
Db	753	CCACCTGTCCGCCGACAGACTGGAAATAGTGGGGGGCGGGGGCGCGCTCGGTGTGGCTGA	694		

Qy	2289	TGGGGCTGAAGGGCGGCCAGAACACGCTGTTC	CGCGTGGCTTCCATCGACGACGTGGTGC	2348
Db	693	AGGGGCTGAAGGGCGGCCAGAACACGCTGTTC	CGCGTGGCTTCCATCGACGACGTGGTGC	634
Qy	2349	GCCTGTTCCTGGCCGAGCTGGGCGGAGAGGAG	CGGACCTGGTGCCTCTTCCTTCCTTCCTTC	2408
Db	633	GCCTGTTCCTGGCCGAGCTGGGCGGAGAGGAG	CGGACCTGGTGCCTCTCTTCCTTCCTTCCTTC	574
Qy	2409	TGGGCTTCGTGGAGCATTTTTCCTGGCTGTCA	ACCGGGTGCATCCCTACCAACGTTC	2468
Db	573	TGGGCTTCGTGGAGCATTTTTCCTGGCTGTCA	ACCGGGTGCATCCCTACCAACGTTC	514
Qy	2469	TCACCTTCCAGCCAGCCCGCCCGCCAGCCGCT	TGGCGGCTCACTTCTCCCGTGG	2528
Db	513	TCACCTTCCAGCCAGCCCGCCCGCCAGCCGCT	TGGCGGCTCACTTCTCCCGTGG	454
Qy	2529	CCGACCTGTCTATCATGCGCGCGCTCTATG	CGCGCTTCAACGCCCGAGATCGGAGCGCGC	2588
Db	453	CCGACCTGTCTATCATGCGCGCGCTCTATG	CGCGCTTCAACGCCCGAGATCGGAGCGCGC	394
Qy	2589	TCGACCTGTCCCTCTATCCTTCGAGAGGGGGT	GTCTCAGCCGCTGAGCTGTGTGAAGAAGG	2648
Db	393	TCGACCTGTCCCTCTATCCTTCGAGAGGGGGT	GTCTCAGCCGCTGAGCTGTGTGAAGAAGG	334
Qy	2649	TCTCCGAGTGCATATGGAACAGCGCTCAGCG	GGTCTCTACTTCAAGGATCGGGCCACATCC	2708
Db	333	TCTCCGAGTGCATATGGAACAGCGCTCAGCG	GGTCTCTACTTCAAGGATCGGGCCACATCC	274
Qy	2709	AGTCCCTCTTCAGTTTCATCAGAGTTTGAGCC	CAGTAGGTGGGAATCTTATCCATGACC	2768
Db	273	AGTCCCTCTTCAGTTTCATCAGAGTTTGAGCC	CAGTAGGTGGGAATCTTATCCATGACC	214
Qy	2769	CACCTTCTCAAAACCCCTCCATGGTTTACAGA	ACCCCTTTTAAAGCTGTAAAGCTTTGTGAG	2828
Db	213	CACCTTCTCAAAACCCCTCCATGGTTTACAGA	ACCCCTTTTAAAGCTGTAAAGCTTTGTGAG	154
Qy	2829	GTTCCGAGGTGTATTTTCTCTTTTTCAGTTT	TGGGAACTGAAGCCCGCAGAGAGGGGAAT	2888
Db	153	GTTCCGAGGTGTATTTTCTCTTTTTCAGTTT	TGGGAACTGAAGCCCGCAGAGAGGGGAAT	94
Qy	2889	GATATGCCAAAGTCACACACGGGATGGCAGG	CTGGAAGCTGAAGCTCGATCACTTGGCTC	2948
Db	93	GATATGCCAAAGTCACACACGGGATGGCAGG	CTGGAAGCTGAAGCTCGATCACTTGGCTC	34
Qy	2949	CAATCATCAACCTCACTCTGCCCTCCAGCA		2981
Db	33	CAATCATCAACCTCACTCTGCCCTCCAGCA		1

RESULT 14
US-09-864-761-22111/C
: Sequence 22111, Application US/09864761
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLES
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864, 761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180, 312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263, 6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236, 359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666

```

RESULT 14
US-09-864-761-22111/c
; Sequence 22111, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```


PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 22111
LENGTH: 639
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC000134.14
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
OTHER INFORMATION: EST_HUMAN HIT: BF526328.1, EVALU0 0.00e+00
OTHER INFORMATION: NT HIT: 093237.1, EVALU0 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: 000255, EVALU0 2.00e-43
US-09-864-761-22111

Query Match 7.0%; Score 639; DB 7; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7422 CAGACTTTTCTGGCCAGGGGAGCATCTCCCATCCCTTCGGTCCGATGGAGTGA 7481
DB 639 CAGACTTTTCTGGCCAGGGGAGCATCTCCCATCCCTTCGGTCCGATGGAGTGA 580
QY 7482 ACCCCCTGGGTGGATGGATGCCAGAGCAGGGTCTCTGGAGTTCAGGCACCTGGCCGGC 7541
DB 579 ACCCCCTGGGTGGATGGATGCCAGAGCAGGGTCTCTGGAGTTCAGGCACCTGGCCGGC 520
QY 7542 AACCTTGTCTACCTTGTCTCTCCCACTGGCCAGGTGGCCAGAGGTGGCCATAGTG 7601
DB 519 AACCTTGTCTACCTTGTCTCTCCCACTGGCCAGGTGGCCAGAGGTGGCCATAGTG 460
QY 7602 AGCCGAGAGGCCAGGCGGCGGAGGAGCCGTGGGGGAGAGCCCGGGAAGGC 7661
DB 459 AGCCGAGAGGCCAGGCGGCGGAGGAGCCGTGGGGGAGAGCCCGGGAAGGC 400
QY 7662 CGGCGGCGGCGGCGGAGGAGTCCAGCCAGAGAGCCCGGCGGCGGCGGAGAGCCA 7721
DB 399 CGGCGGCGGCGGCGGAGGAGTCCAGCCAGAGAGCCCGGCGGCGGCGGAGAGCCA 340
QY 7722 GCACTGGACAAGGGCTTGGGCGACCGGCCAGGGTGCAGTGTACAGACCCCGCCGAGGCCT 7781
|||||

DB 339 GCACTGGACAAGGGCTTGGGCGACCCGGCCAGGGTGCAGTGTACAGACCCCGCCGAAAGCCT 280
QY 7782 CTTGGGACTGTCTGGCAGCAGCCCGAGCCCTGAAGTGGGAGCAGCGCTCAGTGGCCA 7841
DB 279 CTTGGGACTGTCTGGCAGCAGCCCGAGCCCTGAAGTGGGAGCAGCGCTCAGTGGCCA 220
QY 7842 GCACCCGCGCAGCATCACACCCGCGGAGGGTCCAGTGTCTCACTTTCAGAGTGAAGATG 7901
DB 219 GCACCCGCGCAGCATCACACCCGCGGAGGGTCCAGTGTCTCACTTTCAGAGTGAAGATG 160
QY 7902 AAGGGCATGAAGAGTGTCTGGTGGCCACCAAGATCAATCGAGCGCCATCAAGCTGCAA 7961
DB 159 AAGGGCATGAAGAGTGTCTGGTGGCCACCAAGATCAATCGAGCGCCATCAAGCTGCAA 100
QY 7962 CTCACGGCAGCTGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTAC 8021
DB 99 CTCACGGCAGCTGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTAC 40
QY 8022 ACTCTCTCTTCTCTCAAGCGGCGCAGCGCAAGGCGCTCTGA 8060
DB 39 ACTCTCTCTTCTCTCAAGCGGCGCAGCGCAAGGCGCTCTGA 1
RESULT 15
US-09-948-941-1028
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH CANCER, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL000788
CURRENT APPLICATION NUMBER: US/09/948,941
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/231,328
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 12618
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1028
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-948-941-1028
Query Match 6.5%; Score 600.6; DB 5; Length 601;
Best Local Similarity 99.8%; Pred. No. 1.9e-05;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6964 GCCTTCAGTTTTCATGGCCAGACTCCCTCCCTCAGCTGAGGGTGGAGGTAGGTATG 7023
DB 1 gccttcagtttctcatggccagactccctccctcagctgaggggtaggtaggtac 60
QY 7024 GTCTCGCTATGATGGCTTTTATAAGAAAGAGAGTTCTAAGATGTTCCCAACCTAT 7083
DB 61 gtctcgctatgattggtttataaaagaaagaggttctaagaatgttcccaacat 120
QY 7084 GCTTACCTTTTTCGAGCCAGCGGTCTTTCCTCAGTGGGGGCGCTGGCCCTGTGCCCTCT 7143
DB 121 gcttaccttttcgagccagcggttcttcctcaggtgctcctcaggtggtcctctgacctct 180
QY 7144 GCTAAGGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTTCCTCCCTCCCTCCAGGCCA 7203
DB 181 gctaaagggtagtaagagactgattgtgacctcctcctcctcctcctcctcctcctcctc 240
QY 7204 GAGCCAAAGGTTCGCCCTCCAGGACCTTGTCTTGTCCGCCACTGCTGCTGCTGCTTACGA 7263
DB 241 gaggcaaggttcgccctcctcaggaacctgagtgcttcgccccacctgctgctgcttctacga 300
QY 7264 CGGCATCTGCAATGGAGGAGGCGCAGTCCACGCCCTGTCTGCACCTGGGCTGGGCCAC 7323
DB 301 yggcatctgcaaatggag 360
QY 7324 CTTTCTGTGCACTCCCTAGGCGCTTTTGGAGGAGAGGTGAGGAGCAGCTGCACAGAGGT 7383
|||||

```
|||||
Db 361 ctttctgtgcagtccttagcgcttttgaggagacaggtgaggacagctgcacagaggt 420
QY 7384 CTGGGCACTACAGGTGGTGACAGCACGGGCTTGTGACACTTTTCTGCCCCAGGGGC 7443
Db 421 ctgggcactacaggtggtgacagcagccacgggctgtcagactttctgcccaggggc 480
QY 7444 AGCATCTGCCCATCCCTTTCGGTGCCGATGGGACTGAGACCCCTGGGTGGGATGGGATG 7503
Db 481 agcatctgccatcccttcggtgccgatggactggagacccctgggtgggatgggatg 540
QY 7504 GCCAGAGCAGGGTCTCTGAGTTCAGGCACCTGCGCGCAACCTTGCTCTCACCTTGCTCT 7563
Db 541 gccagagcaggggtcctggagttccagccactggccggcaaccttgctctcaccttgctct 600
QY 7564 C 7564
Db 601 c 601
```

Search completed: October 14, 2001, 08:34:02
Job time: 17247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 07:39:50 ; Search time 13115.4 seconds
(without alignments)
10342.904 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGCTCTGTAACCTCTGCC.....AGCCACGGGGCGCCGCCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues
Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098000_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098001_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098002_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098003_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098004_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098005_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US098006_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US098007_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US098008_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US098009_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US098010_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US098011_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1425.2	15.5	2770	22	US-09-577-408-136
2	1365.2	14.9	2121	22	US-09-577-408-6141
3	1239.4	13.5	2840	49	US-60-172-360-22297
4	1006.8	11.0	32768	53	US-60-213-178-196
5	749.4	8.2	2196	12	US-08-865-337-2
6	746.4	8.1	812	12	US-08-865-337-5
7	706.2	7.7	784	29	US-09-726-211-1570
8	639	7.0	639	1	PCT-US01-00663-18748
9	600.2	6.5	632	49	US-60-177-646-792
10	522.2	5.7	603	20	US-09-539-800-12611
11	500	5.3	513	51	US-60-196-710-1772
12	486.2	5.3	551	21	US-09-540-229-133603
13	480.2	5.2	510	49	US-60-177-646-29
14	471.4	5.1	538	12	US-08-865-337-3
15	452	4.9	464	1	PCT-US01-00663-5551
16	417.2	4.5	549	27	US-09-698-013-349
17	412.2	4.5	4227	22	US-09-577-408-7026
18	402	4.4	443	28	US-09-716-990-1013
19	393	4.3	405	17	US-09-332-782-18044
20	393	4.3	405	19	US-09-515-694-18044
21	392.8	4.3	102818	54	US-60-226-178-2393
22	392.8	4.3	102818	55	US-60-233-468-2393
23	389.8	4.2	24167	1	PCT-US01-01334-8463
24	387.8	4.2	227949	54	US-60-226-178-2425
25	387.8	4.2	227949	55	US-60-233-468-2425
26	387.6	4.2	32768	53	US-60-213-162-64
27	387.4	4.2	46829	56	US-60-245-222-40
28	386	4.2	32768	53	US-60-212-655-53
29	383.8	4.2	32768	53	US-60-212-350-4
30	383.8	4.2	32768	53	US-60-213-248-12
31	383.2	4.2	32148	1	PCT-US01-01339-6906
32	382.6	4.2	30716	55	US-60-230-435-46
33	380.6	4.1	73052	56	US-60-243-780-4
34	379	4.1	32768	53	US-60-213-178-119
35	378.6	4.1	32768	57	US-60-229-515-290
36	376.4	4.1	100448	54	US-60-254-168-48
37	375.8	4.1	20009	52	US-60-206-046-15
38	373.4	4.1	32768	54	US-60-229-514-27
39	372.2	4.1	30996	53	US-60-212-656-136
40	371.2	4.0	190532	53	US-60-212-664-240
41	370.8	4.0	83750	56	US-60-245-201-48
42	369.8	4.0	86859	56	US-60-248-798-36
43	369.6	4.0	174127	53	US-60-212-664-282
44	369.2	4.0	8018	1	PCT-US01-01324-4917
45	369.2	4.0	8018	1	PCT-US01-01329-3322

ALIGNMENTS

```

RESULT
US-09-577-408-136
; Sequence 136, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinks, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/577,408
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 8502
; SOFTWARE: pt_gct_genes Version 1.0
; SEQ ID NO 136
; LENGTH: 2770
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (979)...(1033)
; OTHER INFORMATION: similar to g182507 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-136

Query Match 15.58; Score 1425.2; DB 22; Length 2770;
Best Local Similarity 99.18; Pred. No. 1.5e-103;
Matches 1433; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2288 ATGGGGCTGAAGCGCGCCAGAGAGCGGTGTCCCGCTGCCTCCATCGACGACGTGGTG 2347
Db 1 atggggctgaagcgccagagagcggtgtcccgctgcctccatcgacgacgtggtg 60

Qy 2348 CGCCTGTTGTGCGGAGCTGGCGGAGGAGCGGACCTGGTGGTCTCTTCTTCTGGTG 2407
Db 61 cgcctgttgtgctggagctggcgagagagcgacgtgtcccgctgcctccatcgacgacgtggtg 120

Qy 2408 CTGGGCTTCTGGGAGCATTTCTGTGCTGCAACCGCGTCATCCCTACCAAGCTTCCCGAG 2467
Db 121 ctgggcttctgggagcatttctgtgctgcaaccgcgctcatccctacacagcttcccgag 180

Qy 2468 CTCACCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2527
Db 181 ctcaaccttccagcgccgagcgccgagcgccgagcgccgagcgccgagcgccgagcgccgag 240

Qy 2528 GCCGACCTGTCTATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2587
Db 241 gccgacctgtctatcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 300

Qy 2588 GTCGACCTGTCTATCTCTGAGAGGGGTGTCTCCAGCGGTGAGCTGTGTAAGAAG 2647
Db 301 gtcgacctgtctatctctgagaggggtgtctccagcggtgagctgtgtaagaag 360

Qy 2648 GTCTCCGATGATATGAAAGAGCGCTACGCGCTCTCTACTTCAAGGATCGGGCGCGCATC 2707
Db 361 gtctccgatgataatgaaagagcgctacgcgctctctacttcaaggatcgggcgccacatc 420

Qy 2708 CAGTCCCTCTTCAGCTTTCATCATGAGGTTGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2767
Db 421 cagtccctcttcagcttctcatcacaggttggagcgcgagtaggtggagatctctatccatgac 480

Qy 2768 CCACCTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAAAGCTTGTGA 2827
Db 481 ccacttcttcaaaacccctccatggtttacagaaaccccttttaagaactgtaaagcttga 540

Qy 2828 GGTTCGGCAGGTGTTATTTCTCTCTTTGCGAGTTGGGAAACTGAACCCCGAGAGGGGAAA 2887
```

```

Db 541 ggttcgagcaggtgtattttctctctttgagtggtggaactggaagccagagagaggggaaa 600
Qy 2888 TGATATGCCAAAGTCACACACGGGATGCGAGGCTTGGAGGTGAAGCCTGATCATTGGCT 2947
Db 601 tgaatgcccgaagtcacacacgagcatggcagggctggaagtcgaagcctgaacacttggct 660
Qy 2948 CCAATCATCAACCTCAGCTCTGCCCTCAGCAGCCCGACCCCTTGGCCACTGAACAGCTA 3007
Db 661 ccaatcatcaacctcaacctctgccccctcagcacccttggccactgaacagcta 720
Qy 3008 CAGGAGTTCTAAGCATGAGACACAGAGGGCGGCGACGATTTAGGGGCGACAGAGATGAA 3067
Db 721 caggagttctaaagcatgagacacagagggcgagcagatttagggggcaagaagatgaa 780
Qy 3068 ATTGGGCTGCATTTGAGGCGCTTAAACAAATAATGGCTATGAGATTTTATTTTATTTT 3127
Db 781 attgggctgcatttgaggcagcttaaacaaaaataatggctaagaagatttttttttttt 840
Qy 3128 TTTTCTGAGACAGGGTCTCACTCTGTCCCGCAGGCTGGAGTGCAGTGTGTGATCATGG 3187
Db 841 ttttttgagacaggggtctcaactctgtccccaggtgagtgagtgatgacagag 900
Qy 3188 CTCAGTCAGGCTCAGTCTCCCTGGGCTCAGAGATCTCCCAACCTCAGCCCTCCTGAGTAG 3247
Db 901 ctcaactgcagctcagctctcctgggtcagagatactcccaacctcagcctcctgagtag 960
Qy 3248 CTGAGAGTACAGGATGACGCTGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3307
Db 961 ctgagagtcacagggcagctggctggctggtggttaatttttttttttttttttttttt 1020
Qy 3308 GGTGTCTCACTATGTGGCGCCAGAGCTGGTGTGAACCTCTGGGCTCAAGTGATCTGCCGCG 3367
Db 1021 ggtgtcactatgtggccagactggtctgaactcttgggtcgaagtgtatctgcccgc 1080
Qy 3368 CTCAGTCTCCAAATGCTGGGATTTACAGGTGTGAGCCACCGCAACTGTGTGGCTATGAAA 3427
Db 1081 ctcaactctcccaaatgctgggtattacaggtgtgagccaccgcaactggtggcctatgaaa 1140
Qy 3428 ATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3487
Db 1141 atttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 1200
Qy 3488 AATCTCGGCTCACTCAAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3547
Db 1201 aatctcggctcaactgcaagctctgctctctctctctctctctctctctctctctctct 1260
Qy 3548 CAGCCTCTCTGAGTAGCTGGGACTACAGGAGGCTGCCACCATGCCCTGCTGCTGCTGCTGCT 3607
Db 1261 cagcctctctgagtagctgggactacaggtgcccgcacccgcccggcgtctctctctct 1320
Qy 3608 GGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3667
Db 1321 gatttttttagtagagacaggggtttccacctgttagcagaggtatgctcgatctctctgacc 1380
Qy 3668 TCGTATCCCGCGCGCTTGGCTCTCCAAAGTCTCGGGATTACAGCGCTGAGCCACCGCAC 3727
Db 1381 tctgtatcccgcgctcggctcctccaaagtgtctgggtattacagggcgtgagccaccgccc 1440
Qy 3728 CTGGTC 3733
Db 1441 ccggcc 1446
```

```

RESULT
US-09-577-408-6141/c
; Sequence 6141, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinks, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
```

; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792
; CURRENT APPLICATION NUMBER: US/09/577,408
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 8502
; SOFTWARE: pt_gct_genes Version 1.0
; SEQ ID NO 6141
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1003)...(1386)
; OTHER INFORMATION: similar to g13002527 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-6141

Query Match 14.9%; Score 1365.2; DB 22; Length 2121;
Best Local Similarity 98.1%; Pred. No. 8.1e-99;
Matches 1413; Conservative 0; Mismatches 23; Indels 4; Gaps 3;
QY 2288 ATGGGGCTGAAGCGCGCCAGAGAGAGCTGTTCCCGCTGCGCTCCATCGAGAGGTGGTG 2347
DB 2121 ATGGGGCTGAAGCGCGCCAGAGAGAGCTGTTCCCGCTGCGCTCCATCGAGAGGTGGTG 2062
QY 2348 CGCTGTTTCTCGCGAGCTGGCGGAGAGAGCGGACCTGGTCTCTTCCCTTGGTG 2407
DB 2061 CGCTGTTTCTCGCGAGCTGGCGGAGAGAGCGGACCTGGTCTCTTCCCTTGGTG 2002
QY 2408 CTGGGCTTCGTGAGCATTTTCTGGCTGTCAACCGCGTCTATCCCTACCAAGTTCGCGAG 2467
DB 2001 CTGGGCTTCGTGAGCATTTTCTGGCTGTCAACCGCGTCTATCCCTACCAAGTTCGCGAG 1942
QY 2468 CTACCTTCAGCCAGCCCGCCCGCGCGCTGGCGGCTCCTACCTTCTCCCGTG 2527
DB 1941 CTACCTTCAGCCAGCCCGCCCGCGCGCTGGCGGCTCCTACCTTCTCCCGTG 1882
QY 2528 GCGACCTGTCTATCTGCGCGCTCTATGCGCGCTTACCGCCAGATCGAGGCGCC 2587
DB 1881 GCGACCTGTCTATCTGCGCGCTCTATGCGCGCTTACCGCCAGATCGAGGCGCC 1822
QY 2588 GTGACCTGTCCCTCTATCTCGAGAGAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAG 2647
DB 1821 GTGACCTGTCCCTCTATCTCGAGAGAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAG 1762
QY 2648 GTCTCCGATGTCATATGGAACAGCCTCAGCGCTCTCTACTTCAAGGATCGGGCCACATC 2707
DB 1761 GTCTCCGATGTCATATGGAACAGCCTCAGCGCTCTCTACTTCAAGGATCGGGCCACATC 1702
QY 2708 CAGTCCCTCTTACGCTTCATCAGAGTTGGAGCCAGTGTGGGAATCTTATCCATGAC 2767
DB 1701 CAGTCCCTCTTACGCTTCATCAGAGTTGGAGCCAGTGTGGGAATCTTATCCATGAC 1642
QY 2768 CCACTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAAAGCTTGTGA 2827
DB 1641 CCACTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAAAGCTTGTGA 1582
QY 2828 GGTTCGCGAGGTGTATTTCTCTTTTGCAGTTGGGAAACTGAAGCCAGAGGGGAAA 2887
DB 1581 GGTTCGCGAGGTGTATTTCTCTTTTGCAGTTGGGAAACTGAAGCCAGAGGGGAAA 1522
QY 2888 TGATATGCCAAGTTCACACAGGCGATGCGGGCTGGAAGTGAAGCCTGATCACTTGGCT 2947
DB 1521 TGATATGCCAAGTTCACACAGGCGATGCGGGCTGGAAGTGAAGCCTGATCACTTGGCT 1462
QY 2948 CCAATCATCAACCTCACCTCTGCCCCCTCAGCACCCGCCCTTTCGCACTGAACAGCTA 3007
DB 1461 CCAATCATCAACCTCACCTCTGCCCCCTCAGCACCCGCCCTTTCGCACTGAACAGCTA 1402
QY 3008 CAGGAGTTCCTAAGCATGAGACACAGAGGGCGGACAGAGATTTAGGGGGCAGAGATGAA 3067
|||||

DB 1401 CAGGAGTTCCTAAGCATGAGACACAGAGGGCGGACAGATTTAGGGGGCAGAGATGAA 1342
QY 3068 ATTGGGCTGCATTTGAGGCGAGTTAAACAAATATATGCTATGATGAAGATTTTTTTTTTTT 3127
DB 1341 ATTGGGCTGCATTTGAGGCGAGTTAAACAAATATATGCTATGATGAAGATTTTTTTTTTTT 1282
QY 3128 TTTTGTGAGACAGGGTCTCAGTCTGTGCCCCAGGCTGGAGTGCAGTGGTGTGATCATGG 3187
DB 1281 TTTTGTGAGACAGGGTCTCAGTCTGTGCCCCAGGCTGGAGTGCAGTGGTGTGATCATGG 1222
QY 3188 CTCAGTGCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACTCAGCCTCCTCGAGTAG 3247
DB 1221 CTCAGTGCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACTCAGCCTCCTCGAGTAG 1162
QY 3248 CTGAGATACAGGATGCACCCGCTGGTG-CTGGTTAAATTTTGTATTTTGTATTTTCTACAGA 3306
DB 1161 TTGGGATTACAGGACCTGCCACCATGCTGGCTAAATTTTGTATTTTGTATTTTGTACAGA 1102
QY 3307 TGGTGTCTCACTATGTGGGCCAGAGCTGGTCTTGAACCTCTTTGGGCTCAAGTGAATCTGCCG 3366
DB 1101 TGGTGTCTCACTATGTGGGCCAGAGCTGGTCTTGAACCTCTTTGGGCTCAAGTGAATCTGCCG 1042
QY 3367 CCTCAGTCTCCAAATGCTGGGATTACAGTGTGAGCCACCGCAACTGGTGGCCTATGAA 3426
DB 1041 CCTCAGTCTCCAAATGCTGGGATTACAGTGTGAGCCACCGCAACTGGTGGCCTATGAA 982
QY 3427 AATTTTTTTTTTTTTCAGACGGCTCTCACTCTGTCGCCAGGCTGGAGTGCAGTGGTG 3486
DB 981 AATTTTTTTTTTTTTCAGACGGCTCTCACTCTGTCGCCAGGCTGGAGTGCAGTGGTG 922
QY 3487 CAATCTGGGCTCACTGCAAGCTCTGCTCTCTGCTTTCATGCCATTTCTCTGCTCTGCTGCC 3546
DB 921 CAATCTGGGCTCACTGCAAGCTCTGCTCTCTGCTTTCATGCCATTTCTCTGCTCTGCTGCC 862
QY 3547 TCAGCCTCTCAGTGTGGGACTACAGAGCCCTGCCACCATGCTCTGGCTAAATTTTTTTT 3606
DB 861 TCAGCCTCTCAGTGTGGGACTACAGAGCCCTGCCACCATGCTCTGGCTAAATTTTTTTT 802
QY 3607 TGGATTTTATAGTAGAGAC-GAGGTTTCCACCATGTTAGCCAGGATGGTCTCCTCTCTGA 3665
DB 801 TGGATTTTATAGTAGAGAGGGGTTTCCACCATGTTGGCCAGGCTGGTCTCGAACTCTCTGA 742
QY 3666 CCTC--GTGATTCGCCCGCTTGGCCTCCCAAAAGTGTGGGATTACAGGCTGAGCCACC 3723
DB 741 CCTCAGGTGATCCGCCCGCTTGGCCTCCAGAGTGTGGGATTACAGGCTGAGTGAAGC 682
RESULT 3
US-60-172-360-22297
; Sequence 22297, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 22297
; LENGTH: 2840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2836
; OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-22297

	Query Match	11.0%	Score 1006.8	DB 53	Length 32768	
	Best Local Similarity	99.2%	Pred. No. 3e-71			
	Matches 1074	Conservative 0	Mismatches 3	Indels 6	Gaps 6	
Qy	8034	CTCAAGCGCGAGCGCAAGGCCTCTGAACCTACTTGGGAGCTTCGGACCGCTTCTGGGGACC	8093			
Db	32768	CTCAAGCGCGAGCGCAAGGCCTCTGAACCTACTTGGGAGCTTCGGACCGCTTCTGGGGACC	32709			
Qy	8094	CAGGCTCGG-CCTTAGTCCCCCAACTCTGAGGCCATGTTCTGCCCGCCAGCCCCAAGGGGA	8152			
Db	32708	CAGGCTCGGCGCTTAGTCCCCCAACTCTGAGGCCATGTTCTGCCCGCCAGCCCCAAGGGGA	32649			
Qy	8153	CAGGCCTCACCTCTACCCAAACCCCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACC	8212			
Db	32648	CAGGCCTCACCTCTACCCAAACCCCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACC	32589			
Qy	8213	CACGATTTTCTCAGCTCAGAACCCAGGGCTCTGCCCGCAGTCGTTAGAAATATAGTCTCTCT	8272			
Db	32588	CACGATTTTCTCAGCTCAGAACCCAGGGCTCTGCCCGCAGTCGTTAGAAATATAGTCTCTCT	32529			

Db 1801 CCTTAGTCCCCCAACTGTGAGCCCATGTTCTGCCCCCAGCCCCAAAGGGGAGAGGCGCTCA 1860
Qy 8162 CTTCTACCAAAACCTAGTTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 8221
Db 1861 CTTCTACCAAAACCTAGTTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 1920
Qy 8222 CTCAGAGCTAGAAACCCAGGGGTCTGCCCCAGTCTGTTAGATATATAGGTCTCTTCTCCAGA 8281
Db 1921 CTCAGAGCTAGAAACCCAGGGGTCTGCCCCAGTCTGTTAGATATATAGGTCTCTTCTCCAGA 1980
Qy 8282 ATCCAGCCGGCCCAATGAAACCTCAGCTGGGTCTTAATTACAGTCTTTAAAGGCCCA 8341
Db 1981 ATCCAGCCGGCCCAATGAAACCTCAGCTGGGTCTTAATTACAGTCTTTAAAGGCCCA 2040
Qy 8342 GCCCTAGAAACCCAAAGCTCTCTCTGGG-AACCGCTCACCTAGAGCCAGACCAACGTTAC 8400
Db 2041 GCCCTAGAAACCCAAAGCTCTCTCTGGGNAACCGTCTACCTAGAGCCAGACCAACGTTAY 2100
Qy 8401 TCAGGGTCTCCAGAGCTGTAGGAGCTGAGGTTCACCCCTTAACCCAAAGGAGCACAGG 8460
Db 2101 TCAGGGTCTCCAGAGCTGTAGGAGCTGAGGTTCACCCCTTAACCCAAAGGAGGAAANGG 2160
Qy 8461 TCCACCTCCAGCCGGG 8478
Db 2161 TCCACCTCCAGCCGGG 2178

RESULT 6

US-08-865-337-5
; Sequence 5, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 812 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT02

; CLONE: 1313372
US-08-865-337-5
Query Match 8.1%; Score 746.4; DB 12; Length 812;
Best Local Similarity 97.2%; Pred. No. 2.7e-50;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;
Qy 7687 CCAAGCCAGAGGAGCCCGCCGAGAGCCAGACTGGACAAGGCGCTTGGGCACCG 7746
Db 1 CCAAGCCAGAGGAGCCCGCCGAGAGCCAGACTGGACAAGGCGCTTGGGCACCG 60
Qy 7747 GCCAGGGTGCAGTGTCCAGGACCCCGCCGAGAGCCCTCTGGGACTGTCTGCTGGCAGACCC 7806
Db 61 GCCAGGGTGCAGTGTCCAGGACCCCGCCGAGAGCCCTCTGGGACTGTCTGCTGGCAGACCC 120
Qy 7807 GAGGCCCTGAGGTTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCACACCCGCGG 7866
Db 121 GAGGCCCTGAGGTTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCACACCCGCGG 180
Qy 7867 AGGTTCAGTGTCTCACTTTCCAGAGTGAAGAGATGAAGGGCATGAAGGAGCTGTCTGGTGG 7926
Db 181 AGGTTCAGTGTCTCACTTTCCAGAGTGAAGAGATGAAGGGCATGAAGGAGCTGTCTGGTGG 240
Qy 7927 CCACCAAGATCAACTCGAGCGCCATCAAGTGCACACTACGCGACAGTGCAGAGTGCAGA 7986
Db 241 CCACCAAGATCAACTCGAGCGCCATCAAGTGCACACTACGCGACAGTGCAGAGTGCAGA 300
Qy 7987 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTCTTTCTCAAGCGGCGAGC 8046
Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTCTTTCTCAAGCGGCGAGC 360
Qy 8047 GCAAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCACCGCTCCG-CCT 8105
Db 361 CAAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCACCGCTCCGCGCT 420
Qy 8106 TAGTCCCGCCAACTCTGAGCCCATGTCTGCCCGCCAGCCAAAGGGGAGACGGCTCACCTC 8165
Db 421 TAGTCCCGCCAACTCTGAGCCCATGTCTGCCCGCCAGCCAAAGGGGAGACGGCTCACCTC 480
Qy 8166 TACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 8225
Db 481 TACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 540
Qy 8226 AGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGAATATAGGTCTCTTCTCCAGAAATCC 8285
Db 541 AGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGAATATAGGTCTCTTCTCCAGAAATCC 600
Qy 8286 CAGCGGCCCAATGGAACCTCAGCTGGGTCTTAATTACAGTCTTTAAAGGCCCGAGGCC 8345
Db 601 CAGCGGCCCAATGGAACCTCAGCTGGGTCTTAATTACAGTCTTTAAAGGCCCGAGGCC 660
Qy 8346 CTAGAAACCCAAAGCTCTCTCCCTCGG-AACCGCTCAGCTAGAGCCAGACCAAGCTTACTCAG 8404
Db 661 CTAGAAACCCAAAGCTCTCTCCCTCGGNAACCGTTCAGCTAGAGCCAGACCAAGCTTAYTCAG 720
Qy 8405 GGCTCTCTCCAGTTGTAGGAGCTGAGGTTCACCTTTAACCCCAAGGGGAGACGACGCTCC 8464
Db 721 GGCTCTCTCCAGTTGTAGGAGCTGAGGTTCACCTTTAACCCCAAGGGGAGGAAANGGTCC 780
Qy 8465 ACCTCCAGCCCGGG 8478
Db 781 ANCTTCAGCCCGGG 794

RESULT 7

US-09-726-211-1570
; Sequence 1570, Application US/09726211
; GENERAL INFORMATION:
; APPLICANT: Hunter, John J.
; APPLICANT: Shvjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; THEREFOR

Qy 7782 CCTGGAGCTGTGGTGACAGCCGAGGCGCTGAAGGTGCGACAGCGCTCAGGTGCA 7841
Db 279 CCTGGAGCTGTGGTGACAGCCGAGGCGCTGAAGGTGCGACAGCGCTCAGGTGCA 220
Qy 7842 GCACCCGAGCATCACCACCGCGAGGCTCAGTGTCTCACTTTCCAGAGTGAGAAGTG 7901
Db 219 GCACCCGAGCATCACCACCGCGAGGCTCAGTGTCTCACTTTCCAGAGTGAGAAGTG 160
Qy 7902 AAGGCGATGAAGGAGCTGCTGTGGCCACCAAGATCAACTCGAGCGGCATCAAGCTGCAA 7961
Db 159 AAGGCGATGAAGGAGCTGCTGTGGCCACCAAGATCAACTCGAGCGGCATCAAGCTGCAA 100
Qy 7962 CTCAGGCGACAGTCGCAAGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGACTAC 8021
Db 99 CTCAGGCGACAGTCGCAAGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGACTAC 40
Qy 8022 ACTCTGTCTTTCTTCAAGCGCGCAGCGCAAGGCGCTCTGA 8060
Db 39 ACTCTGTCTTTCTTCAAGCGCGCAGCGCAAGGCGCTCTGA 1

RESULT 9
US-60-177-646-792/c
; Sequence 792, Application US/60177646
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000210
; CURRENT APPLICATION NUMBER: US/60/177,646
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 4226
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 792
; LENGTH: 632
; TYPE: DNA
; ORGANISM: HUMAN
US-60-177-646-792

Query Match 6.5%; Score 500.2; DB 49; Length 632;
Best Local Similarity 97.8%; Pred. No. 8e-39;
Matches 619; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 4732 AAGGATCATACATCGCTGTGACCGCAAGATGGAGGTGGCGTTTCATGTGTGTCATCA 4791
Db 632 AAGGATCATACATCGCTGTGACCGCAAGATGGAGGTGGCGTTTCATGTGTGTCATCA 573
Qy 4792 ACCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGCAGGTGAGGG 4851
Db 572 ACCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGCAGGTGAGGG 513
Qy 4852 CTGAGCCAAATGGGAGGACTGGGCTAGGCCAGACTGTGCTGTGGGACCTGGGCA 4911
Db 512 CTGAGCCAAATGGGAGGACTGGGCTAGGCCAGACTGTGCTGTGGGACCTGGGCA 453
Qy 4912 GGGCAGCTTCCCTTCCTGAGCTTCAGCTTCCCTCCCTGGAAATGGTTAGTAATTC 4971
Db 452 GGGCAGCTTCCCTTCCTGAGCTTCAGCTTCCCTCCCTGGAAATGGTTAGTAATTC 393
Qy 4972 TGGCTGGCCCTTTCCCGAGGCTTTGGGAGAGTAGAATTGAGATGCTGAAATGCTTTGAC 5031
Db 392 TGGCTGGCCCTTTCCCGAGGCTTTGGGAGAGTAGAATTGAGATGCTGAAATGCTTTGAC 333
Qy 5032 TCCATTAAGGGCTGGTCCCGAGATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTTGG 5091
Db 332 TCCATTAAGGGCTGGTCCCGAGATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTTGG 273
Qy 5092 TTTGACCCCGACCTCTGCGCGATAGGCTAAGGACCGCTTCTCCTGTTCCGTGGCT 5151
Db 272 TTTGACCCCGACCTCTGCGCGATAGGCTAAGGACCGCTTCTCCTGTTCCGTGGCT 213

Qy 5152 CATAACTCTCTCTCCCTTCCGCTCCTAGAAGCTGCTGTGGCTGCTCTATGACCTGGGACATCT 5211
Db 212 CATAACTCTCTCTCCCTTCCGCTCCTAGAAGCTGCTGTGGCTGCTCTATGACCTGGGACATCT 153
Qy 5212 GGAAGGTCAGTAGAGGAGTGGCCAGGCTGCGCCTGGTGGAGCCGGGGGCTGGGTGG 5271
Db 152 GGAAGGTCAGTAGAGGAGTGGCCAGGCTGCGCCTGGTGGAGCCGGGGGCTGGGTGG 93
Qy 5272 CAGCTCTGAATTATGATCTCTTCCCTAGGTACCCCTTGGCTTAGGGAACCTGGCAGATCTA 5331
Db 92 CAGCTCTGAATTATGATCTCTTCCCTAGGTACCCCTTGGCTTAGGGAACCTGGCAG-TCTA 34
Qy 5332 GAGGAGCTGGAGCCACCCCTGGCCGCGCCAGAC 5364
Db 33 GAGGAGCTGGAGCCACCCCTGGCCGCGCCAGAC 1

RESULT 10
US-09-539-800-12611
; Sequence 12611, Application US/09539800
; GENERAL INFORMATION:
; APPLICANT: Sellhame, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 12611
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01347785
US-09-539-800-12611

Query Match 5.7%; Score 522.2; DB 20; Length 603;
Best Local Similarity 98.1%; Pred. No. 1e-32;
Matches 560; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Qy	7547	TGCTCTCACCTTGCTCTCCCCACTGGCCCGAGGTGGCGCAGAGTGGCCATAGTAGAGCCG	7606
Db	388	TGCTCTCACCTTGCTCTCCCCACTGGCCCGAGGTGGCGCAGAGTGGCCATAGTAGAGCCG	329
Qy	7607	AGAGGCCAGGGCGCCGAGGCCGAGAGCCGTGGGCGAGGAGGCCCGGGAGAGCCCGGCG	7666
Db	328	AGAGGCCAGGGCGCCGAGGCCGAGAGCCGTGGGCGAGGAGGCCCGGGAGAGCCCGGCG	265
Qy	7657	CGGGGGCCACCGCGGAGTCCAAAGCCAGAGAGCCCCCGCCGCCAAGAGCCAGCACT	7726
Db	268	CGGGGGCCACCGGTGGAGTCCAAAGCCAGAGAGCCCCCGCCGCCAAGAGCCAGCACT	209
Qy	7727	GGACAAGGGCGTGGGCACCGGCCAGGGTGCAGTGTGAGAGCCCCCGCGAAGCCCTCTGG	7786
Db	208	GGACAAGGGCGTGGGCACCGGCCAGGGTGCAGTGTGAGAGCCCCCGCGAAGCCCTCTGG	149
Qy	7787	GACTGTGCGTGGCACGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACC	7846
Db	148	GACTGTGCGTGGCACGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACC	89
Qy	7847	CGCAGCATCACCCCGCGGAGGGTCCAGTGTCTACTTTCCAGAGTGAAGAAGTGAAGGG	7906
Db	88	CGCAGCATCACCCCGCGGAGGGTCCAGTGTCTACTTTCCAGAGTGAAGAAGTGAAGGG	29
Qy	7907	CATGAAGAGCTGCTGGTGCCACCACAG	7934
Db	28	CATGAAGAGCTGCTGGTGCCACCACAG	1

RESULT 12
US-09-540-229-133603
; Sequence 133603, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.

```

, FILE REFERENCE: PD-1033 C1P
, CURRENT APPLICATION NUMBER: US/09/540,229
, CURRENT FILING DATE: 2000-03-31
, Prior application data removed - refer to file wrapper
, NUMBER OF SEQ ID NOS: 193582
, SOFTWARE: PERL Program
, SEQ ID NO 133603
, LENGTH: 551
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, NAME/KEY: misc_feature
, OTHER INFORMATION: Incyte ID No: hu01071364
, NAME/KEY: unsure
, LOCATION: 389, 423, 466, 484
, OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-133603

```

		Matches	543;	Conservative	0;	Mismatches	7;	Indels	5;	Gaps
Qy	8224	CCAGCTCAGAACCAGGGGCTCTGCCCGACTCGTTAGAAATATAGGTCTCTTCTCCAGAAAT	8283							
Db	1	ccagctcagaaccoca-ggcctctcccagtcgtagaatataggctctctctccccgaat	59							
Qy	8284	CCCAGCGGCCCAATTGAAAACCTCACGCTGGTGCTCAATAATTCACAGTCTTTAAAGGCCACGC	8343							
Db	60	cccagcgcgccaatcgaaaacctcacctgggtgctcaattaccacgactctttaaagcgccacg	119							
Qy	8344	CCCTAGAAACCCAAGCTCTCTTCGGAACCGCTCACCTAGAGCCAGACCAACGTTTACTTCA	8403							

[illegible]

RESULT 14

; GENERAL INFORMATION:

APPLICANT: COVITZ, Peter
APPLICANT: Tang, Y TOM

;
; TITLE OF INVENTION: HUMAN

; NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

STREET: 3174 Porter

; STATE: CA

21F: 34304
: COMPUTER READAB

; COMPUTER: IBM Compat

```

; SOFTWARE: FASLSEQ IOI W
;
; CURRENT APPLICATION DATA:

```

;
; FILING DATE: Herewith

; PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:

; ATTORNEY/AGENT INFORM

REGISTRATION NUMBER: 36,
REFERENCE/DOCKET NUMBER:

TELEPHONE: 415-855-0555

TELEX: INFORMATION FOR COUNTRIES

; LENGTH: 538 base pairs

; STRANDEDNESS: single

LIBRARY: BRAINOT14

US-08-865-337-3

Query Match

Best Local Similarity 98.38; Pred. No. 1e-28;

QY 2251 GGCTTGCTTTCAGCGCGCCGCCACCGCCGCCGCGGCTGAGGCGGCGCCAGAA 2310
Db 20 GCTAGTGTGGGATGCGCGCGCCGCCACCGCCGCCGCGGCTGAGGCGGCGCCAGAA 79
QY 2311 GAGCTGTTCCTCGCTGCTCCATCGACGAGGTGGTGGGCTGTTCCTCGCGAGCTGGG 2370
Db 80 GAGCTGTTCCTCGCTGCTCCATCGACGAGGTGGTGGGCTGTTCCTCGCGAGCTGGG 139
QY 2371 CCGAGAGGAGCGGACCTGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2430
Db 140 CCGAGAGGAGCGGACCTGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 199
QY 2431 GGCTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCGCCCGC 2490
Db 200 GGCTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCGCCCGC 259
QY 2491 CCGGACCGCGCTGGGCGCTCACCTACTTTCGGTGGCGAGCTGCTATCATCGCGC 2550
Db 260 CCGGACCGCGCTGGGCGCTCACCTACTTTCGGTGGCGAGCTGCTATCATCGCGC 319
QY 2551 CCTCTATGCCCGCTTACCGCGCCAGATCCGAGGCGCGCTGCGACTGTCCCTCTATCCTCG 2610
Db 320 CCTCTATGCCCGCTTACCGCGCCAGATCCGAGGCGCGCTGCGACTGTCCCTCTATCCTCG 379
QY 2611 AGAAGGGGGTGTCTCCAGCGCTGAGTGGTGAAGAAGGTCTCCGAGTGCATATGGAACAG 2670
Db 380 AGAAGGGGGTGTCTCCAGCGCTGAGTGGTGAAGAAGGTCTCCGAGTGCATATGGAACAG 439
QY 2671 CCTCAGCGCTCTACTTCAAGATCGGCGCCACATCCAGTCCCTCTTTCAGCTTCATCAC 2730
Db 440 CCTCAGCGCTCTACTTCAAGATCGGCGCCACATCCAGTCCCTCTTTCAGCTTCATCAC 499
QY 2731 AGG 2733
Db 500 AGG 502

RESULT 15
PCT-US01-00663-5551/c
; Sequence 5551, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: EB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 5551
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000134.14

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
PCT-US01-00663-5551
Query Match 4.9%; Score 452; DB 1; Length 464;
Best Local Similarity 99.8%; Pred. No. 3.5e-27;
Matches 463; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 7714 AGAAGCCAGCACTGGCAAGGGCTGGGCACCGCCAGGGTGCAGTGTACAGACCCGCC 7773
Db 464 AGAAGCCAGCACTGGCAAGGGCTGGGCACCGCCAGGGTGCAGTGTACAGACCCGCC 405
QY 7774 GGAAGCCCTCTGGGACTGTCTGGCACAGCCCGAGGCCCTGAAGGTGGCAGCAGGCTC 7833
Db 404 GGAAGCCCTCTGGGACTGTCTGGCACAGCCCGAGGCCCTGAAGGTGGCAGCAGGCTC 345
QY 7834 AGGTGCCAGCACCCGCGAGCATCACCGCGGGAGGGTCCAGTGTCTACTTTCAGAGTG 7893
Db 344 AGGTGCCAGCACCCGCGAGCATCACCGCGGGAGGGTCCAGTGTCTACTTTCAGAGTG 285
QY 7894 AGAAGATGAAGGGCATGAAGGAGCTGTGTGGCCACCAAGATCAACTCAGCGCCATCA 7953
Db 284 AGAAGATGAAGGGCATGAAGGAGCTGTGTGGCCACCAAGATCAACTCAGCGCCATCA 225
QY 7954 AGCTGCAACTCACGGCACAGTTCGCAAGTGCAGATGAAGAAGCAGAAAAGTGTCCACCCCTA 8013
Db 224 AGCTGCAACTCACGGCACAGTTCGCAAGTGCAGATGAAGAAGCAGAAAAGTGTCCACCCCTA 165
QY 8014 GTGACTACACTCTGTCTTCTCAAGCGGCAGCGCAAGGCTCTGAACTACTGGGACT 8073
Db 164 GTGACTACACTCTGTCTTCTCAAGCGGCAGCGCAAGGCTCTGAACTACTGGGACT 105
QY 8074 TCGGACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTC 8132
Db 104 TCGGACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTC 45
QY 8133 TGCCCCCAGCCCCAAAGGGGAGAGCGCTCACCTCTACCCAAACCC 8176
Db 44 TGCCCCCAGCCCCAAAGGGGAGAGCGCTCACCTCTACCCAAACCC 1

Search completed: October 14, 2001, 07:44:28
Job time: 37994 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 03:51:57 ; Search time 218.79 Seconds
(without alignments)
7943.124 Million cell updates/sec

Title: US-09-380-337-3

Perfect score: 9180

Sequence: 1 CTGCTCTGAACCTCTGCCC.....AGCCACGGGGCGCGCCGCCG 9180

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	749.4	8.2	2196	2	US-08-865-337A-2
2	746.4	8.1	812	2	US-08-865-337A-5
3	471.4	5.1	538	2	US-08-865-337A-3
4	348	3.8	14796	4	US-08-975-080-35
5	347.6	3.8	31571	1	US-08-323-443B-1
6	347.6	3.8	53526	3	US-08-658-136-2
7	347.6	3.8	53577	3	US-08-658-136-1
8	342.4	3.7	14796	4	US-08-975-080-35
9	329.8	3.6	4421	5	PCT-US95-07201-9
10	329.8	3.6	4421	5	PCT-US95-07201-9
11	325.2	3.5	17949	4	US-08-273-411-2
12	324.8	3.5	3373	1	US-08-273-411-2
13	324.8	3.5	8174	1	US-07-914-281-5
14	324.8	3.5	8174	1	US-08-393-246-5
15	324.8	3.5	8174	1	US-08-525-058A-5
16	324.8	3.5	8174	2	US-08-696-731-5
17	324.8	3.5	8174	4	US-09-042-531-5
18	324.8	3.5	8174	5	PCT-US91-00899-3
19	320	3.5	14636	4	US-09-173-914-6
20	318.4	3.5	11613	1	US-08-484-044-10
21	318.2	3.5	5543	2	US-08-687-080-101
22	315.2	3.4	35060	3	US-08-814-095-7
23	304.4	3.3	4803	4	US-09-197-636-1
24	304.4	3.3	4803	4	US-09-197-636-3
25	303.2	3.3	3286	4	US-09-211-417-2
26	302.2	3.3	7620	1	US-07-767-135-1
27	302.2	3.3	7620	1	US-07-841-652-1

c 28	295.6	3.2	10380	4	US-09-077-354B-3	Sequence 3, Appli
c 29	294.8	3.2	14636	4	US-09-173-914-6	Sequence 6, Appli
c 30	293.2	3.2	1442	2	US-08-454-557C-120	Sequence 120, App
c 31	293.2	3.2	1442	2	US-08-340-426D-120	Sequence 120, App
c 32	293.2	3.2	1442	2	US-08-450-673C-120	Sequence 120, App
c 33	292.8	3.2	4421	2	US-08-257-963B-9	Sequence 9, Appli
c 34	292.8	3.2	4421	5	PCT-US95-07201-9	Sequence 9, Appli
c 35	291.8	3.2	3035	1	US-08-726-725-2	Sequence 2, Appli
c 36	290.4	3.2	22481	5	PCT-US95-07201-43	Sequence 43, Appli
c 37	290.4	3.2	35060	3	US-08-814-095-7	Sequence 7, Appli
c 38	290	3.2	4803	4	US-09-197-636-1	Sequence 1, Appli
c 39	290	3.2	4803	4	US-09-197-636-3	Sequence 3, Appli
c 40	289.2	3.2	72928	3	US-09-009-913-1	Sequence 1, Appli
c 41	288.2	3.1	3035	1	US-08-726-725-2	Sequence 2, Appli
c 42	288.2	3.1	6769	1	US-08-480-784-20	Sequence 20, Appli
c 43	288.2	3.1	6769	1	US-08-483-553-20	Sequence 20, Appli
c 44	288.2	3.1	6769	1	US-08-487-002-20	Sequence 20, Appli
c 45	288.2	3.1	6769	1	US-08-483-554B-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-08-865-337A-2
; Sequence 2, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
; US-08-865-337A-2

```
Query Match      8.2%  Score 749.4;  DB 2;  Length 2196;
Best Local Similarity 97.1%  Pred. No. 5.2e-148;
Matches 775;  Conservative 4;  Mismatches 17;  Indels 2;  Gaps 2;

Qy 7683 GAGTCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCAGCAGTGGACAAAGGGCGCTGGC 7742
Db 1381 GAGNCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCAGCAGTGGACAAAGGGCGCTGGC 1440

Qy 7743 ACCGGCCAGGGTGCAGTCTCAGGACCCCGCCCGAGGCTCTCGGAGCTGCTGGCACA 7802
Db 1441 ACCGGCCAGGGCGGAGTCTCAGGACCCCGCCCGAGGCTCTCGGAGCTGCTGGCACA 1500

Qy 7803 GCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCCACG 7862
Db 1501 GCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCCACG 1560

Qy 7863 CCGAGGGTCCAGTCTCAGTCTTCCAGAGTGAAGATGAAGGCATGAAGAGCTGCTG 7922
Db 1561 CCGAGGGTCCAGTCTCAGTCTTCCAGAGTGAAGATGAAGGCATGAAGAGCTGCTG 1620

Qy 7923 GTGGCCACCAAGATCACTCAGCGCCATCAAGCTGCAACTCACGGCACAGTCCCAAGTG 7982
Db 1621 GTGGCCACCAAGATCACTCAGCGCCATCAAGCTGCAACTCACGGCACAGTCCCAAGTG 1680

Qy 7983 CAGATGAAGAAAGCAAGTGTCCACCCCTAGTGACTACACTCTCTCTTCTTCTCAAGCG 8042
Db 1681 CAGATGAAGAAAGCAAGTGTCCACCCCTAGTGACTACACTCTCTCTTCTTCTCAAGCG 1740

Qy 8043 CAGCGCAAGGCCCTCTGAACACTGCGGAGCTTCGGAGCGCTTGTGGGACCCAGGCTCCG 8102
Db 1741 CAGCGCAAGGCCCTCTGAACACTGCGGAGCTTCGGAGCGCTTGTGGGACCCAGGCTCCG 1800

Qy 8103 -CCTTAGTCCCGCAACTCTGAGCCATGTTCTGCGCCAGCCAGCCAAAGGGACAGCGCTCA 8161
Db 1801 CCTTAGTCCCGCAACTCTGAGCCATGTTCTGCGCCAGCCAGCCAAAGGGACAGCGCTCA 1860

Qy 8162 CCTCTACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 8221
Db 1861 CCTCTACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 1920

Qy 8222 CTCAGCTCAGAAACCCAGGCTCTGCCCCAGTCTGAGATATAGTCTCTTCTCCAGCA 8281
Db 1921 CTCAGCTCAGAAACCCAGGCTCTGCCCCAGTCTGAGATATAGTCTCTTCTCCAGCA 1980

Qy 8282 ATCCAGCGCGCCCAATGAACACTCAGCGTGGGTCTTAATTTACAGTCTTTAAAGGCCCA 8341
Db 1981 ATCCAGCGCGCCCAATGAACACTCAGCGTGGGTCTTAATTTACAGTCTTTAAAGGCCCA 2040

Qy 8342 GCCCTTAGAAACCAAGCTCTCTCTCGG -AACCGCTCACCTAGAGCAGACCAACGTTAC 8400
Db 2041 GCCCTTAGAAACCAAGCTCTCTCTCGGNAACCGTACCTAGAGCAGACCAACGTTAY 2100

Qy 8401 TCAGGGCTCTCCAGCTTGTAGAGCTGAGGTTTCACCCCTTAACCCAAAGGAGCACAG 8460
Db 2101 TCAGGGCTCTTCCAGCTTGTAGAGCTGAGGTTTACCCCTTACCCAAAGGAGGANGG 2160

Qy 8461 TCCCACTTCCAGCCCGGG 8478
Db 2161 TCCCACTTCCAGCCCGGG 2178
```

RESULT 2
US-08-865-337A-5
; Sequence 5, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murty, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN

```
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337A
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT02
CLONE: 1313372
US-08-865-337A-5
```

```
Query Match      8.1%  Score 746.4;  DB 2;  Length 812;
Best Local Similarity 97.2%  Pred. No. 1.5e-147;
Matches 772;  Conservative 4;  Mismatches 16;  Indels 2;  Gaps 2;

Qy 7687 CCAAGCCAGAGAGAGCCCGCCCGCCCAAGAGCAGCAGTGGACAAAGGGCGCTGGGACCG 7746
Db 1 CCAAGCCAGAGAGAGCCCGCCCGCCCAAGAGCAGCAGTGGACAAAGGGCGCTGGGACCG 60

Qy 7747 GCCAGGTGTCAGTCTCAGGACCCCGCCCGGAGGCTCTCTGGGACTGTCTGGGACAGGCC 7806
Db 61 GCCAGGTGTCAGTCTCAGGACCCCGCCCGGAGGCTCTCTGGGACTGTCTGGGACAGGCC 120

Qy 7807 GAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCAGCCGAGCATCACCCAGCGCGG 7866
Db 121 GAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCAGCCGAGCATCACCCAGCGCGG 180

Qy 7867 AGGTCTCAGTCTCAGTCTTCCAGAGTGAAGAGTGAAGGAGTGAAGGAGTGAAGGAGTGA 7926
Db 181 AGGTCTCAGTCTCAGTCTTCCAGAGTGAAGAGTGAAGGAGTGAAGGAGTGAAGGAGTGA 240

Qy 7927 CCACCAAGATCAACTCGAGGCCATCAAGTGCACCTCAGGTCAGGTCAGGTCAGGTCAGG 7986
Db 241 CCACCAAGATCAACTCGAGGCCATCAAGTGCACCTCAGGTCAGGTCAGGTCAGGTCAGG 300

Qy 7987 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTCTTCTCTCAAGCGGACG 8046
Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTCTTCTCTCAAGCGGACG 360

Qy 8047 GCAAGGCCCTCTGAACACTCTGGGACTTTCGACCGGCTTGTGGGACCCAGGCTCCG-CCT 8105
Db 361 CAAAGGCCCTCTGAACACTCTGGGACTTTCGACCGGCTTGTGGGACCCAGGCTCCGCGCT 420

Qy 8106 TAGTCCCGCAACTCTGAGCCCATGTTCTGCGCCCGCCAGCCCAAGGGGACAGGCCCTCACCTC 8165
```



```
; FILING DATE: 20-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-975-080-35

Query Match      3.8%; Score 348; DB 4; Length 14796;
Best Local Similarity 66.8%; Pred. No. 1.4e-63;
Matches 624; Conservative 0; Mismatches 285; Indels 25; Gaps 8;

Qy 382 GCCAGGCACAGTGGCTCACACCTGTAAATCCAGCAGCTTTTGGAGGCCGAGGTAGAGGAT 441
Db 6289 GCCGGGTGTGTGGCTCACACCTGTAAATCCAGCAGCTTTGGAGGCCGAGGTAGAGGAT 6230
Qy 442 CACTTGAGGTGAGGAGTTCGAGAGCAGCTGACCAACATGCTGACACCCCTGTCTCTACT 501
Db 6229 CAC--GAGTCAAGGATCAAGACCCCTCTGGCTAACAGGTGAACCC--GTCTCTACT 6173
Qy 502 AAGTAAATACAAAATPAGCCAGTGTGTGGCAGGACCTGTAATCCCGGCTACTTTGG 561
Db 6172 AA-AAATACAAAATPAGCCGGGTGTGTGGCGGGCTGTAGTCCAGCTACTCGG 6114
Qy 562 GAAGCTGAGGAGAGATCACTTGAACCTGGAGGAGAGGTTGCGAGTGGCGGAGTC 621
Db 6113 GAGGTGAGGAGGAGAGATGCTTGAACCTGGGCGGGAGCTTGCAGTGGCGGAGAT 6054
Qy 622 ACCCAGCTGACCTCAGCCTGAGTGACAGAGCAGACTCTCTCTCAAAACAA----- 673
Db 6053 CCGTGGCTGACCTTAGCTGGCGGAGAGAGCAAGACTCCATCTCAATAATAATAAA 5994
Qy 674 ---ATAACAAATACACTACTCTTTGGCGGGTAAGTGGTTCAGCGCTGTAAATTTAGCA 730
Db 5993 TTTAAATAATAAAATACACTTCTGAGCGGAGGTGTGTGGCTACACCTATATATCTAGCA 5934
Qy 731 CTTTGGGAGGTGAGGCGGAGATCACTTCAAGTGTAGGGTTCGAGACCACTCTGGCCA 790
Db 5933 CTTTGGGAGGTGAGGCGGAGATCACTTCAAGTGTAGGGTTCGAGACCACTCTGGCCA 5874
Qy 791 ACATGGTGAACCCCATCTCTACTTAAATACAAAAGTTTTCCTGGTGTGGTGGCGGAC 850
Db 5873 ACATGGTGAACCCCTCTCTACCAAAATACAAA--TTACCTGGGTGTGGTGGCAT 5815
Qy 851 GCCTAATATCCAGTACTTGGGACTTTTATTAAGCGGAATCTCACTCTGTGTGCCAG 910
Db 5814 GCCTGTAACTCTGGCTACTTGGAGGCT-----GAAGCAGGCAAAATCACTTGAAGCCA 5762
Qy 911 GCTGGAGTGCAGTGCAGATTTCTGGCTCACTGAAGCTCCCGCTCCAGGTTCAAGGG 970
Db 5761 GGAGCGGAGGTTGCAGTGTATCCGAGATCAGGCCACCCCTACTCCAGCTGGGTAAACAG 5702
Qy 971 ATTCCCGCGCTCAGCTCCCAAGTAGCTGGGAATCCCTGTCTCTGCAAAAAA 1030
Db 5701 TGAACCTGTGTCTAGACAAAATAAATAAAGGAGGCTGTGGTTCACCCCTGTAAATC 5642
Qy 1031 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1090
Db 5641 CCAGCAGCTTTGGGAGGCAAGTGGCGCAATCAGAGTCAAGAGTTCAAGACCAAGCCTG 5582
Qy 1091 TTATATGATATATATATATATATATATATATATATATATATATATATATATATATAT 1150
Db 5581 GGCAATGTTGGTGAACCCCATCTCTACTTAAATAAATAAATAAATAAATAAATAAATAAATA 5523
```

```
Qy 1151 GCAGCGCTGTGATCACAGTACTCGGAGGCTGAGGCACGAGAAATCGCTTGGCCCGTGA 1210
Db 5522 ACTTCCCTGTAAATCCAGTACTTGGGAGGCTGAGGCAGGAGAACTACTTGAACCCGAGA 5463
Qy 1211 AGTCGAGGCTGACATGAGCCAGATCGAGCCACTGCGCTATTCAGGCTGGCTGA--AAGAGAA 1269
Db 5462 GGCAGAGGTTGACGTGAGCCAAAGATTGCACCATTTGCACCTCCAGCTGGCAACAAGAGTG 5403
Qy 1270 AGACCGTGTCTCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1303
Db 5402 AAACCTATCTCAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5369
```

```
RESULT 5
US-08-323-443B-1
; Sequence 1, Application US/08323443B
; Patent No. 5654170
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W.
; APPLICANT: LANDES, GREGORY M.
; APPLICANT: BURN, TIMOTHY C.
; APPLICANT: CONNORS, TIMOTHY D.
; APPLICANT: DACKOWSKI, WILLIAM R.
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,443B
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, S. Peter
; REGISTRATION NUMBER: 25,351
; REFERENCE/DOCKET NUMBER: 0372/0A462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: PKD1 GENOMIC
; US-08-323-443B-1
```

```
Query Match      3.8%; Score 347.6; DB 1; Length 31571;
Best Local Similarity 78.7%; Pred. No. 2.2e-63;
Matches 496; Conservative 0; Mismatches 114; Indels 20; Gaps 6;

Qy 3115 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3174
Db 6512 TCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 6571
```


APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 53577 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match 3.8%; Score 347.6; DB 3; Length 53577;
Best Local Similarity 78.7%; Pred. No. 2.8e-63;
Matches 496; Conservative 0; Mismatches 114; Indels 20; Gaps 6;
QY 3115 TTTTCTTTTGTGAGACAGGCTCTCACTGTCTCCCGCAGGCTGGAGTGCAGT 3174
Db 7496 TCTTTTGTGTTGTTGAGAGGAGTCTGGCTGTCTCAACCCAGGCTGGAGGACAAT 7555
QY 3175 GGTGTGATCATGGCTCACTGCAGGCTCAGTCTCCCTGGGCTCAGAGATCCTCCACACCTCA 3234
Db 7556 GGTGTGATCTCGGCTCACTGCAGGCTCCTCCCGGTTCAAGCGATCTCCTCGCTCA 7615
QY 3235 GCCTCTGAGTAGCTGAGAGTACAGCATGCAC--CGTGGTGTCTGTTAATTTTGTAT 3292
Db 7616 GCCTCTGAGTAGCTGGATGAGAGCGCGGCCACACAGCGCCGCTAATTTTAAAAAT 7675
QY 3293 TTTTCTGAGAGATGGTCTCTCACTATGTGGCCAGACAGTGGTCTTGAACCTTTGGGCTC 3352
Db 7676 ATTTTGTAGAGATGGGCTTTTCCACCATGTTGGTCAAGCTGGTCTTGAACCTTTGGCCTT 7735
QY 3353 AGTGTATGCGCCGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTGAG--CCACC 3407
Db 7736 AGTGTATGCTGTGCTCGGCTCCCAAAGTGTCTGGGATTACAGGTGTGAGTGTAT 7795
QY 3408 GCAACTGGTGGCTATGAAATTTTTTTTTTTTTCAGACGGCGTCTCACTCTGTCGCGCC 3467
Db 7796 TTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7855
QY 3468 AGGCTGGAGTGCAGTGGTGCATCTCGGCTCACTCCAAAGCTCTGCCTCTGCTTTTCATGC 3527
Db 7856 AGGCTGGAGTGCAGAGTGCATCTCAGTCACTGCAAGCTCCGCTCTGCTGGTTTCACGC 7915
QY 3528 CATTTCTCTGCTCTGCTCAGCTCCTGAGTACTGGGACTACAGGAGCTGGCCACA 3587
Db 3528 CATTTCTCTGCTCTGCTCAGCTCCTGAGTACTGGGACTACAGGAGCTGGCCACA 3587

Db 7916 CATT-----CTCTGCTCAGCCTCTCTGAGTAGCTGGACT---GGTGGCCGTCACCA 7965
QY 3588 TCCCTGGCTAAATTTTGTGATTTTGTAGTAGACAGAGTTTCAACCATGTTAGCCAGG 3647
Db 7966 TGCCAGCTAA--TTTGTGATTTTGTAGTAGACAGGGGTTTCAACCTGTTAGCCAGG 8022
QY 3648 ATGGTCTCGATCTCTGACCTCGTGATCCGCCCGCTTGGCTTCCCAAAGTGTGGGAT 3707
Db 8023 ATGGTCTCGATCTCTGACCTCGTGATCTCCGCTCAGCCTCCCAAAGTGTGGGAT 8082
QY 3708 ACAGGCTGAGCCAGCCAGCTGGTCAAAA 3737
Db 8083 ACAGGCTTGAAGCCAGCCCTGTCTTTTAAA 8112
RESULT 8
US-08-975-080-35
Sequence 35, Application US/08975080
Patent No. 6245523
GENERAL INFORMATION:
APPLICANT: Altieri, Dario C.
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
CELLULAR APOPTOSIS, AND ITS MODULATION
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,080
FILING DATE: 20-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,435
FILING DATE: 20-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 14796 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match 3.7%; Score 342.4; DB 4; Length 14796;
Best Local Similarity 76.0%; Pred. No. 2e-62;
Matches 497; Conservative 0; Mismatches 131; Indels 26; Gaps 5;
QY 3118 TTTTCTTTTGTGAGACAGGCTCTCACTGTCTCCCGCAGGCTGGAGTGCAGTGT 3177
Db 5671 TTTTCTTTTGTGAGACAGGCTCTCACTGTCTCCCGCAGGCTGGAGTGCAGTGT 5730
QY 3178 GTGATCATGGCTCACTGCAGGCTCAGTCTCCCTGGGCTCAGAGATCCTCCAAACCTCAGCC 3237
Db 5731 CTGATCTCGGATCACTGCAACCTCCGCTCCTGGGCTCAAGTGAATTTGCTTTCAGCC 5790
QY 3238 TCCCTGAGTAGTGAAGTACAGGATGACCGTGTGCTGTTTAAATTTTGTATTTT 3297
Db 3238 TCCCTGAGTAGTGAAGTACAGGATGACCGTGTGCTGTTTAAATTTTGTATTTT 3297

Db 5791 TCCCACTAGCGAGATTACAGGATGTCACACACACCAGGTAATTTTGTGA--TTTT 5848
QY 3298 TTGTAGAGATGGTGTCTCACTATGTGGCCAGACTGTCTTGAACCTTGGGCTCAAGTG 3357
Db 5849 TGGTAGAGACGAGGTTTCCACCATGTGGCCAGGCTGTTTGAACCTCTGACCTCAGGTG 5908
QY 3358 ATCTGCCGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGAGCCAGGCAACTGG-- 3415
Db 5909 ATCCACCCGCTCAGCCTCCCAAAGTCTGAGATTATAGGTGAGCCACACACCTGGC 5968
QY 3416 -----TGGCCTATGAAATTTTTCAGACGCGCTCTCACTCTGT 3462
Db 5969 CTCAGGAAGTATTTTATTTTAAATTTATTTATTTAGATGAGTCTGCTCTGT 6028
QY 3463 CGCCAGGCTGGAGTGGTGTGCAATCTCGGCTCACTCAAGCTCTGCTCTGCTTT 3522
Db 6029 CGCCAGGCTAGAGTCAGCAGCGGGATCTCGGCTCACTCAAGCTCCGCCCCAGGTT 6088
QY 3523 CATGCCATTCCTGCTCTGCTCAGCTCTGCTGAGTCTGAGTCTGAGGACTACAGGAGCTGC 3582
Db 6089 CAAGCCATTT-----CTCCTGCTCAGCTTCCCGAGTAGCTGGGACTACAGGCGCCGC 6141
QY 3583 CACCATCGCTGGCTAAATTTTGTGATTTTGTAGTAGAGACGAGGTTTACACCATGTAG 3642
Db 6142 CACACACCGCGTAA--TTTTTTGTATTTTGTAGTAGAGACGAGGTTTACACCGTGTAG 6199
QY 3643 CCAGGATGGTCTCGATCTCCTGACCTCGTGTATCGCCCGCTTGGCCTCCCAAAGTCTG 3702
Db 6200 CCAGGAGGCTTGTATCTCTGACCTCGTGTATCTGCTGCTGCTGCTGCTCCCAAAGTCTG 6259
QY 3703 GGATTACAGCGGTGAGCCAGCGACCTGGTCAAAATGTTTGTAGACAGAGAAGG 3756
Db 6260 GGATTACAGGTGTAGCCACACACACCGCGCTATTTTATTTTGTAGACAGGG 6313

RESULT 9

US-08-257-963B-9
; Sequence 9, Application US/08257963B
; Patent No. 5840686
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, S.
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEPT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4421 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: J7101
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.1 kb Bam HI
; OTHER INFORMATION: fragment Derived from human placental
; OTHER INFORMATION: genomic DNA
; US-08-257-963B-9

Query Match 3.6%; Score 329.8; DB 2; Length 4421;
Best Local Similarity 77.4%; Pred. No. 5.6e-60;
Matches 484; Conservative 0; Mismatches 117; Indels 24; Gaps 6;
QY 3117 TTTTCTTTTCTTTTGTGACAGAGGTCTCACTCTGTCCCCAGGCTGGAGTGCAGTGG 3176
Db 3491 TATTTTCTTTTCTTTTGTGATGGAGTCTGGCTCTGTACCCAGGCTGGAGTGCAGTGG 3550
QY 3177 TGTGATCATGGCTACACGAGGCTCACTCTCCCTGGGCTCAGAGATCTCCAACTCAGC 3236
Db 3551 CGGATCTCGGCTCAGTGTCAAGCTCTTCTCCCGGGTTTCACTCCCTGCTCTCAAC 3610
QY 3237 CTCTGATAGTCTGAGAGTACAGGCTACAGGCTGGTGGTGTAAATTTTGTATTTT 3296
Db 3611 CTCCGAGTGTGTAATACAGGCTGCGCCACACCGCGACTAATTTTGTATTTT 3670
QY 3297 TTGTAGAGATGGTCTCACTATGTGGCCAGACTGGTCTTGAACCTCTTGGGCTCAAGT 3356
Db 3671 TTAGTAGAGACGGGGTTTCACTGGTGTAGCAGGATGGTCTGGATCTCTGACCTC--GT 3728
QY 3357 GATCTGCCGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTAGCCAGGCTGGG 3415
Db 3729 GATCTGCTGCTCGGCTCCCAAAGTGTGGGATTACAGGCTTGAAGCTTGAAGCTTGAAG 3788
QY 3416 TGGCTATGAAATTTTCTTTTTCAGACGCGCTCTCACTCTGTCCGCCAGGCTGGA 3475
Db 3789 CCTC-----TTATTTTCTTTTGTAGATGGAGTCTCACTGTGACCTGGGCTGGA 3839
QY 3476 GTGCAGTGGTCAATCTCGGCTCACTGCAAGCTCTGCTCTCTTTCATGCCAATCTCC 3535
Db 3840 GTGCAGTGGAGCGATCTCGGCTCACTGCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3895
QY 3536 TGCCTCTGCTCAGCTCTCTGAGTCTAGTCTAGGAGCTTACAGGAGCTGCGCACCCTGCG 3595
Db 3896 ---CTCTGCTCAGCTCTCCCAAGTAGTGGGATTACAGGTTGAGGCTTGAAGCTTGAAG 3952
QY 3596 TAATTTTCTTTTGTAGTAGACGAGGTTTCACTATGTTAGCAGGATGGTCTC 3655
Db 3953 TA---GTTTCTTTTGTAGTAAGATGGGTTTCACTATGTTGGCCAGGCTGCTT 4009
QY 3656 GATCTCTGACCTC--GTGATCCCGCCGCTTGGCTCTCCCAAGTGTGGGATTACAGGC 3713
Db 4010 GAACTCTGACATCAGGTGATCCGCCACCTTAGCCTCCCAAGTGTGGGATTACAGGC 4069
QY 3714 GTGAGCCACCGACCTGTCACAAA 3738
Db 4070 GTGAGCCACCATACCTGGCCAGCA 4094

RESULT 10
PCT-US95-07201-9
; Sequence 9, Application PC/TUS9507201
; GENERAL INFORMATION:

QY	3297	TTTGTAGAGATGGTGTCTCACTATGTGCCCCAGACTGGTCTTGAACATCTTTGGGCTCAAGT	3350
Db	3671	TTAGTAGACACGGGGTTTCAACCGTGTAGCCAGGATGCTCTGGATCTCTTCACTC--GT	3728
QY	3357	GATCTGCCCGCCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTGAGCCACCGCAACTGG	3415
Db	3729	GATCTGCCCTGGCCTGGCCCTCCCAAAGTCTGGGATTACAGGCTTGAAGCCACCGACCCGG	3788
QY	3416	TGCGCTATGAAATATTTTTTTTTTTTTCAGACGGCGTCTCACTCTGTGCGCCCAAGGCTGGA	3475
Db	3789	CTTC-----TTATTTTTTTTTCAGATGGAGTCTCACACTGTCACTGGGCTGGA	3839
QY	3476	GTGCAGTGTGCAATCTCGGCTCACTGCAAGCTCTGGCTCTGTCTGTCTTTCATGCCATCTCC	3535
Db	3840	GTGCAGTGTGAGGAGTCTCGGCTCACTGCAAGCTCTCGGCTCTCGGGTTCAAGAGATTT--	3895
QY	3536	TGCGCTCTCGCTCAGCTCCTGAGTAGTGGGACTACAGGAGCGCTGCCACCATCGCTGGC	3595
Db	3896	---CTCCTCGCTCAGCTCCCAAGTAGCTGGGATTACAGGTGCCACCAACCGCCTGGC	3952
QY	3596	TAATTTTTTTTTTGGATTTTTAGTAGAGACGAGGTTTCAACCATGTTAGCCAGGATGTCTC	3655
Db	3953	TA---GTTTTTTGTATTTTTAGTAAAGATGGGGTTTCAACCATGTTGGCCAGGCTGTCT	4009
QY	3656	GATCTCCTCAGCTC--GTGATCGCGCGCTTTGGCCTCCCAAGTGTCTGGGATTACAGGC	3713
Db	4010	GAACTCTGACATCAGGTGATCCGCCCACTTAGCTCCCAAGTGTCTGGGATTACAGGC	4069
QY	3714	GTGAGCCACCGCACCCTGGTCAAAA	3738
Db	4070	GTGAGCCACCATCTGGCCAGCAA	4094

RESULT 11

US-09-087-465-3/c

; Sequence 3, Application US/09087465A

; Patent No. 6160092

; GENERAL INFORMATION:

; APPLICANT: Vinkemeier, Uwe

; APPLICANT: Chen, Xiaomin

; APPLICANT: Darnell Jr., James E

; APPLICANT: Kuriyan, John

; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS C

; TITLE OF INVENTION: USE

; FILE REFERENCE: 600-1-229

; CURRENT APPLICATION NUMBER: US/09/087,465A

; CURRENT FILING DATE: 1998-05-29

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 3

; LENGTH: 17949

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-087-465-3

	Query Match	3.5%;	Score 325.2;	DB 4;	Length 17949;
	Best Local Similarity	76.2%;	Pred. No. 8.9e-59;		
	Matches 483;	Conservative 0;	Mismatches 133;	Indels 18;	Gaps 6;
Qy	3099	TAATGGCTATCAAGATTTTTTTTTTTTTTTTTTTT	TGAGACAGGGTCTACACTCTGTCTCCC	3158	
Db	3118	TAATCTGTATTTTTTTTTTTTTTTTTTTTTTTT	TAAAGATGACTCTCACACTCTTTTGTGCC	3059	
Qy	3159	CAGGCTGGATGTCAGTGGTG-TGATCATGGCTCACTGCAGCCTCAGTCTCCCTGGGCTCA	3217		
Db	3058	TGGCCGGAGTGAAGTGGAGATGATCTTGGCTCATCTGCAACACCATCCCGCTGGTGTCT	2999		
Qy	3218	GAGATCTCCAACTCAGCCCTCCGTAGTGTAGAGCTACAGGATGACACCGTGGTGGTG	3277		
Db	2998	GTGATTCTCTTGCCCTCAGCCCTCCGTAGTGTAGAGCTACAGGATGACACCGTGGTGGTG	2939		

QY 3278 GTTAATTTTTTGTAT-TTTTTTGTAGAGATGGTGTCTCACTATGTGGCCAGCACTGGTC 3336
Db 2938 AGCTAATTTTTTGTATATTTTTTTAGTAGAGATGGGTTTTTGCCATGTTGCCAGCTGGTC 2879
QY 3337 TTGAACCTCTTGGCTCAAGTATCTGCCCGCTCACTCTCCCAA-TGCTGGGATTACAG 3395
Db 2878 TCAAGTCTGACCTCAGGTGATCCACCTACTTCGGACTCCCAAGTACTAGGATTACAG 2819
QY 3396 GTGTAGAGCCACCGCACTGGTGGCTTATGAAATTTTTTTTTTTTTCAGACGGCTCTC 3455
Db 2818 GCATGAGCC-----ACTGTGGTCCCAACCTGTATTTTTTTTTTCTCAGATGGAGTCTC 2765
QY 3456 ACTCTCTGCCAGGCTGAGTGAGTGTGCAATCTCGGCTCACTGCAAGCTCTCCCTC 3515
Db 2764 GATCTCTTGGCCAGGCTGAGTGAGTGTGATTTTGGCTCACTGCAAGCTCCGCCCTC 2705
QY 3516 CTGCTTTTCATGCCATCTCTCGCTCCTGCTCAGCTCCTGAGTAGCTGGGACTACAGG 3575
Db 2704 CCGGGCTCATGCCATT-----CTCCTGCCCTCAGCTCCCGAGTAGCTGGGACTATAGG 2652
QY 3576 AGCTGCCACCATGCTCGCTAA--TTTTTTTTTGTATTTTAGTAGACGAGGTTTCA 3633
Db 2651 CTCCGCCACCATGCCCGCTGATTTTTTTTTTGTATTTTAGTAGAGATGGGGTTTCA 2592
QY 3634 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTATCGGCCCTTGGCCCTCC 3693
Db 2591 CCGTGTAGCCAGGATGGTCTCGATCTCCTGAACTCGTGACTCGCCTACCTCGGCCCTCC 2532
QY 3694 AAAGTCTGGGATTACAGCGGTGAGCCACCGCAC 3727
Db 2531 AAAGTCTGGGATTACAGCGGTGAGCCGATGCC 2498

RESULT 12

US-08-273-411-2
; Sequence 2, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 104..1201
; OTHER INFORMATION: /note= "Nucleotides 104 through 1201 encode the GDP-L-fuc
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
US-08-273-411-2

Query Match 3.5%; Score 324.8; DB 1; Length 3373;
Best Local Similarity 76.0%; Pred. No. 5.6e-59;
Matches 491; Conservative 0; Mismatches 122; Indels 33; Gaps 6;

QY 3116 TTTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCACTCTGTCCTCCCGAGCTGGAGTGCAGTG 3175
Db 1753 TTTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCGCTCTGTTGCCCGAGCTGGAGTGCAGTG 1812
QY 3176 GTGTGATCATGGCTCACTCGAGCTCAGTCTCCCTGGGCTCAGAGATCTCCCAACCTCAG 3235
Db 1813 GCCTGATCTGGCTCACTCAACTTCGCCCTCTCTGTTCAAGGATCTCTCTCTCTCAG 1872
QY 3236 CTCTCTGATGATGAGTACAGGATGACCGTGGTGTGCTGTTAAATTTTTTTTGTATTTT 3295
Db 1873 CTCTCTGATGATGAGTACAGGATGACCGTGGTGTGCTGTTAAATTTTTTTTGTATTTT 1930
QY 3296 TTTTGTAGAGATGGTCTCACTATGTGCCAGAGCTGGTCTTGAACCTCTTGGGCTCAAG 3355
Db 1931 TTTAGTAGAGACAGGGTTTCCACCATGTTGGCCGGGATGGTCTCGATCTCTCGACCT--TG 1988
QY 3356 TGATCTGCCCGCTCAGTCTCCCAAA--TGCTGGGATTACAGGCTCTGAGCCACCCCAAC-- 3412
Db 1989 TCATCCACCTGTCTTGGCCCTCCCAAGTGTGGGATTACTGGCATGAGCCACTGTGCCCA 2048
QY 3413 -----TGCTGGCCTATGAAATTTTTTTTTTTTTCAGACGGCTCTCAG 3457
Db 2049 GCCCGGATATTTTTTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2108
QY 3458 TCTGTGCCCGAGCTGAGTGCAGTGGTCAATCTCGGCTCACTGCAAGCTCTGCCCTCCT 3517
Db 2109 TCTGTAGCCAGCCAGAGTGCAGTGGCGGATCTCAGCTCACTGCAAGCTCTGCCCTCCC 2168
QY 3518 GCTTTTCATGCCATTTCTCTGCCCTCCTGCCCTCAGCTCCTGAGTAGCTGGGACTACAGGAG 3577
Db 2169 GGGTTTCATGCCATT-----CTGCCCTCAGCTCCTGAGTAGCTGGGACTACAGGCG 2218
QY 3578 CTTGCCACCATGCTCTGCTAAATTTTTTTTTTTTGGATTTTTTAGTAGAGAGAGTTTACCAT 3637
Db 2219 CCGGCCACACGCCCGGCTAA-TTTTTTTTGTATTTTATTTAGTAGAGACGGGTTTTCATCGT 2277
QY 3638 GTTAGCCAGGATGGTCTCGATCTCTGACCTCTGATCCGCCCGCTTGGCCCTCCCAAG 3697
Db 2278 GTTAACCCAGGATGGTCTCGATCTCTGACCTCTGATCTGCCACCTCTGCCCTCCCAAG 2337
QY 3698 TGCTGGGATTACAGGGCTGAGCCACCGCACTGCTGCTCAAAATGTTT 3743
Db 2338 TGCTGGGATTACCGCGCTGAGCCACCATGCTGGCCCGGATAATTT 2383

RESULT 13

US-07-914-281-5
; Sequence 5, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

[illegible]

```
Query Match      3.5%; Score 324.8; DB l; Length 8174;
Best Local Similarity 76.0%; pred. No. 7.9e-59;
Matches 491; Conservative 0; Mismatches 12; Indels 33; Gaps 6
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 04:05:21 ; Search time 644.85 Seconds
(without alignments)
8938.725 Million cell updates/sec

Title: US-09-380-337-3

Perfect score: 9180

Sequence: 1 CTGCTCTTGAACCTCTGCCC.....AGCCACGGCGCGCCGCCCG 9180

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9179	100.0	9180	19	AAV57416 Tumour suppressor
2	1312.2	14.3	2772	19	AAV57415 Tumour suppressor
3	749.4	8.2	2196	20	AAV80659 Human multiple end
4	746.4	8.1	812	20	AAV80662 Human multiple end
5	471.4	5.1	538	20	AAV80660 Human multiple end
6	352.4	3.8	119950	20	AAV90201 Human ves1 gene.
7	348	3.8	14796	19	AAV27941 Survivin gene. Ho
8	347.6	3.8	53526	19	AAV94101 Human PKD1 gene.
9	347.6	3.8	53577	17	AAV18551 Human polycystic k
10	347.6	3.8	53577	19	AAV94108 Human PKD1 locus b
11	345.8	3.8	66566	21	AAA53450 Human thioredoxin

C 12	344.8	3.8	122186	22	AAC89560	Human histone deac
C 13	343.6	3.7	26664	21	AAA60207	Human prostate can
C 14	342.8	3.7	51474	22	AAV97846	Human neuroblastom
C 15	342.4	3.7	14796	19	AAV27941	Survivin gene. Ho
C 16	341.4	3.7	26928	20	AAZ32184	Human prothrombin
C 17	336.6	3.7	122186	22	AAC89560	Human histone deac
C 18	336.2	3.7	39198	22	AAV58067	Human polyamine-mo
C 19	335.4	3.7	1470	14	AAQ47355	Myotonic dystrophy
C 20	335	3.6	11967	22	AAV97863	Human neuroblastom
C 21	331.8	3.6	160552	22	AAQ02697	Human glycosyl sul
C 22	329	3.6	66566	21	AAA53450	Human thioredoxin
C 23	328	3.6	1963	21	AAA51400	Chromosome 16g tum
C 24	326	3.6	18402	21	AAZ93705	CYBRP fatty acid
C 25	325.2	3.5	17949	22	AAV89228	Human signal trans
C 26	325.2	3.5	18647	21	AAV21059	Human low adenosin
C 27	325.2	3.5	18648	21	AAA34937	Human adenosine re
C 28	325.2	3.5	72604	20	AAZ10752	Genomic sequence o
C 29	324.8	3.5	3373	18	AAV76768	Human alpha 1,2 fu
C 30	324.8	3.5	8174	12	AAQ13332	GDP-Fuc:beta-D-gal
C 31	324.8	3.5	8174	18	AAQ16177	Human alpha(1,2)-f
C 32	323.6	3.5	160552	22	AAQ02697	Human glycosyl sul
C 33	323.2	3.5	8174	15	AAQ056908	DNA encoding a gly
C 34	322.6	3.5	32367	19	AAV35820	Human SHOX (short
C 35	321.8	3.5	32367	19	AAV35620	*Human SHOX (short
C 36	321.2	3.5	2918	21	AAV73083	*Human MCL1 gene 5'
C 37	321.2	3.5	3204	21	AAV73082	Human melanocortin
C 38	321	3.5	11967	22	AAV97863	Human neuroblastom
C 39	320.2	3.5	15577	19	AAV35616	SHOX gene prelinin
C 40	320	3.5	14636	22	AAV31864	Human DNA-PKcs cod
C 41	319.2	3.5	121162	21	AAC66548	Human kinesin-like
C 42	318.6	3.5	15071	21	AAZ58411	hnRNP A2 genomic c
C 43	318.2	3.5	5543	18	AAV75284	Nucleotide sequenc
C 44	317.6	3.5	49999	20	AAZ23902	Human LOBO homolog
C 45	316.6	3.4	45546	20	AAV23520	Human kidney amino

ALIGNMENTS

RESULT 1	
AAV57416	
ID	AAV57416 standard; cdna; 9180 BP.
XX	
AC	AAV57416;
XX	
DT	11-JAN-1999 (first entry)
XX	
DE	Tumour suppressor gene MEN1.
XX	
KW	Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;
KW	menin; tumour suppressor gene; cancer; marker; diagnosis;
KW	gene therapy; human; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS
FT	2288..8060
FT	/*tag= a
FT	/note= "contains introns"
FT	exon
FT	1680..1766
FT	/*tag= b
FT	/number= 1
FT	intron
FT	167..2264
FT	/*tag= c
FT	/number= 1
FT	exon
FT	2265..2732
FT	/*tag= d
FT	/number= 2
FT	intron
FT	2733..4296
FT	/*tag= e
FT	/number= 2
FT	exon
FT	4297..4505
FT	/*tag= f

FT intron /number= 3
 FT 4506..4715
 FT /*tag= g
 FT /number= 3
 FT 4716..4844
 FT /*tag= h
 FT /number= 4
 FT 4845..5176
 FT /*tag= i
 FT /number= 4
 FT 5177..5217
 FT /*tag= j
 FT /number= 5
 FT 5218..5297
 FT /*tag= k
 FT /number= 5
 FT 5298..5385
 FT /*tag= l
 FT /number= 6
 FT 5386..6024
 FT /*tag= m
 FT /number= 6
 FT 6025..6161
 FT /*tag= n
 FT /number= 7
 FT 6162..6622
 FT /*tag= o
 FT /number= 7
 FT 6623..6758
 FT /*tag= p
 FT /number= 8
 FT 6759..7195
 FT /*tag= q
 FT /number= 8
 FT 7196..7360
 FT /*tag= r
 FT /number= 9
 FT 7361..7577
 FT /*tag= s
 FT /number= 9
 FT 7578..9180
 FT /*tag= t
 FT /number= 10
 FT 8861..8865
 FT /*tag= e

TTATA_signal

WO9839439-A1.

11-SEP-1998.

04-MAR-1998; 98WO-US04258.

05-MAR-1997; 97US-0040269.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;

Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;

Lubensky IA, Manickam P, Marx SJ, Spiegel AM;

WPI: 1998-506360/43.

P-PSDB; AAW29749.

Protein and gene associated with multiple endocrine neoplasia type 1
 - useful in gene therapy and to diagnose sufferers of, and those
 susceptible to, this condition by detecting protein absence or gene
 mutation(s)

Claim 32; Page 57-62; 75pp; English.

This is the nucleotide sequence of human MEN1, a novel tumour
 suppressor gene which is associated with multiple endocrine
 neoplasia type 1. MEN1 genomic DNA can be obtained from genomic

CC libraires using MEN1-specific probes and PCR primers. A cDNA clone
 CC (see AAW57415) is also provided. MEN1 encodes a 67.5 kDa protein
 CC (see AAW28749), termed menin, that exhibits no apparent similarities
 CC to previously known proteins. The lack of a functional menin
 CC polypeptide, either by absence of the protein, its alteration
 CC and/or associated mutations in the MEN1 gene, have been identified
 CC in individuals with familial multiple endocrine neoplasia type 1
 CC (FMEN1) and suffering from multiple endocrine neoplasia type 1.
 CC The identification of MEN1 provides a new window into the mechanism
 CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
 CC MEN1 associated cancers, and provides preclinical identification of
 CC individuals with the FMEN1 syndrome, i.e. cancer-free individuals
 CC that are at high risk of acquiring MEN1 associated tumours. It
 CC also provides a basis for gene therapy. The MEN1 nucleic acids may
 CC also be used therapeutically to produce menin recombinantly and to
 CC produce transgenic animals useful in research.

XX Sequence 9180 BP; 1936 A; 2604 C; 2612 G; 2027 T; 1 other;

Query Match 100.0%; Score 9179; DB 19; Length 9180;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 9180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCTTGAACCTCCTGCCCTCAAGCAATCCTCTGCTTCAGCTTCCCAAGTGTGTAA 60
 Db 1 ctggctcttgaactcctggcctcaagcaatcctcctcagcttcccaagtggtgaa 60
 Qy 61 TTACAGGCATGAGCCTGGCATGAACTTGACACTATTGAGATATACTGCTCAGGTATTTTG 120
 Db 61 ttacaggcatgagcctggcatgaaacttgacactattgagatactggcaggtatttg 120
 Qy 121 TGAATGTCCTCAACTCTGTTTGGCCAGATGTTTTCATGATTAGAGGAGACTTATAA 180
 Db 121 tggaaatgctccctcaactctgtttgccagatggtttctcattgattagaggagattataa 180
 Qy 181 ATTTTGAGGAAATCCAGAGAGGTGAAGAGGTGAAGTAGAGTGGCCAGCAAAATTAATCTGTTT 240
 Db 181 attttgagaaaatccagagaggtgaagagtggaagtaggagcagaaatttaactctgttt 240
 Qy 241 ATTTACTGCTATATACCGAGTGTCTGGNACTTGGCCCTGCTAGTACCAAAATCTGTTT 300
 Db 241 atttactgctatataccagagtgctggaacttgcccatggaagtagtaccataaaactctgtt 300
 Qy 301 TTTTTCGAATGAATAAGCAATAAATGAGTCACCGCTGGAAATTTAGTATTTTCAAAAGT 360
 Db 301 ttttttgaatgaataagcaataaataaagtgagtgaccgtggaatttagtatttcaaaagt 360
 Qy 361 TTCAAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTATATCCAGACATTT 420
 Db 361 ttcaaaagcgttgttgatacagggccaggcacagtggtccacacctglaatccacagcactt 420
 Qy 421 TGGAGGCGGAGTAGGAGGATCACTTGAGGTCAGAGTTCGAGACCAGCCGACCAACAT 480
 Db 421 tggagggccgaggtgagggaggtacacttgaggtcagaggtccgagaccagcctgacacacat 480
 Qy 481 GGTGACACCCCTGCTCTACTTAAGTAAATAACAAAATTAGCCAAAGTGTGGTGGCAGGCA 540
 Db 481 ggtgacacccctgctctacttaagtaaaatacaaaaaattagccaagtggtggtgagggca 540
 Qy 541 CCTGTATCCCGGCTACTTGGGAAGCTTGAGGACGAGCAAGAAATCAGTGAACCTGGGAGGAG 600
 Db 541 cctgtaatccccggctacttgggaagctgagcagaagaatacacttgaacctgggagggcag 600
 Qy 601 AGTGTGAGTGCAGCGGAGATCACCCCTGCTGACCTCCAGCCTGAGTGACAGAGCAGAGCTC 660
 Db 601 aggttgagtgagccgagatcacccctgctcctccagcctgagtcagagggcagagctc 660
 Qy 661 TGTCTCAAAACAAATAAACAATAACTACTCTTTGGCGGGTAAGGTGGTTCAGCCCTGT 720
 Db 661 tgtctcaaaacaaataaacaataactactctttggccgggtgaagggtgttccagcctgt 720
 Qy 721 AATTTAGCACCTTTGGGAGGCTGAGGGGGCAGATCACTTGAAGTTAGGGTTTCAGACC 780

Db 721 aatttttagcaactttgagaggtgagcgggcagatcacttgaggttaggggttcgagacc 780
QY 781 AGTCTGGCCAAACATGTGAAACCCCATCTCTACTTAAATACAAAAGTTTCTTGGGTGT 840
Db 781 agtctggccaaacatggtcgaaacccctctctacttaaaatacacaaaagtttcttgggtg 840
QY 841 GGTGGCGGACGGCTATAATCCAGCTACTTGGGACATTTTTTAAAGACGAATCTCACATC 900
Db 841 ggtggcgagcgctataaattccagctacttggagactttttttaagacggaatctcactc 900
QY 901 TGTGTCGCCAGGCTGGAGTCGAGTGGCAAGATCTGGCTCACTGAAGCCCTCCGCTCCCAG 960
Db 901 tgttgcccaggtggagtgagtggaagattctggtcactgaagcctccgctccag 960
QY 961 GTTCAAGGGATTTCCCGCGCTCACCTCCCAAGTAGCTGGGAATCCCTGTCTGCAAA 1020
Db 961 gttcaagggtattccccgcgtccagctcccaagtagctgggaatccctgtctctgcaa 1020
QY 1021 AAAAAAATAAAAAAATAAATAATATATATATATATATATATATATATATATATATATAT 1080
Db 1021 aaaaaaataaaaaaataaataataataataataataataataataataataataataata 1080
QY 1081 TGTGTGTGTTAT 1140
Db 1081 tgtgtgtgttatgtatatataatttatgtatatgtatatatatatatatatatatatatgt 1140
QY 1141 GAGTGTGGCGACGGCTGTGTATCAGCTACTCTCGGAGGCTGAGCGACGAGAAATCGCTT 1200
Db 1141 gaagtggcgacgcctgtgatacagactactcggaggtgagggcagagaaatcgctt 1200
QY 1201 GAGCCGTGAAGTCGAGGTGCGAGTGCAGCCAGATCGAGCCACTGATTTCCAGCCTGGGC 1260
Db 1201 gagccgtgaagtgcaggtgcagtgagcccagatcgagccactgcatccagcctgggc 1260
QY 1261 GAAAGAAAGACCGTGTCTAAACAAACAAACAAACAAAGCTACTCTTAGCACGTGTATGA 1320
Db 1261 gaagagaagacgtgtgtcaaaacaaacaaacaaacaaagctactcttagcagtgtaga 1320
QY 1321 GTATCTCGCGGGGGAAGTGGAAACGAGTGTGTACAGACAGATAGGCATCTTTATATGT 1380
Db 1321 gatactcgcgggcggaagtgggaacgagtgctgscacagagtaggacatcttataatgt 1380
QY 1381 TAACAGACACTGATACCCAGCTAAAGCGCTGAACACATTTACTCTGCGAGTGTTTAA 1440
Db 1381 taacagacactgataccagctaaagcgctgaacacattactctctggaagtgtaga 1440
QY 1441 AGTATCTGTTTCTCATATTTGTTTATTTTAAATTTTCTGGATCAAGCAACCTGATC 1500
Db 1441 aagtaactgttttctcatattgttttatttttaatttttctggaataagcaacctgac 1500
QY 1501 TTTTTCCTCATAACTTCCGACCGCCGCTGACAGCAAAACCGCAGAGCTCGCGGACC 1560
Db 1501 ttttctctcataacttgcgcacgcacccctgacagcaaacccggcagaagctcgagacc 1560
QY 1561 TCCACCCGAGTCTCAGGTAGTGCCTCCCGGACTACTATTTTCCAGAAAGCACTGCGGGC 1620
Db 1561 tccaccccgagctcgaggtagtgccccggagactacattttccagaagcactgcgggc 1620
QY 1621 ACCTTCTCTGCTGCTGGCTGAAGGGAGGGCCAAATCCCTGAGTATCTTCGGGAAGGAG 1680
Db 1621 acgcttcttgcctgtcgcgctgaagggaagggccaatccctctgagtatctcggaagag 1680
QY 1681 GTGTCCGAGCGCGACCTAGAGATCCAGAGCCACAGCGCGCGCCCGCCGCCCA 1740
Db 1681 gfgtcgggagcgcgagccttagaga tccccagaagccacagcgagcgccccggcgcca 1740
QY 1741 CTATTTCCAGGCTCTCGGGGGCAGGGGTGGCCAGACTCCACTCCCGCGGGGTAGTGC 1800
Db 1741 ctatttccaggtctcgggcgaggggtggggccagactccacttccccggcgagtgatgc 1800
QY 1801 GACCCTAGGGCGGGACTTCATGTCTCCAGAGGCTCCGGCGCGCTCGCGCGCGTGCCT 1860
Db 1801 ggggaaatgatafgccaaagtccacacgcatggcgaggtctggaaagtgaagcctgatca 1860

Db 1801 gaccctagggcgggagcttcatgtcccagaggtctccggcgcggtgctgcgcgcgggtgacct 1860
QY 1861 AGTGTGGGATGTAAAGCGGAGGTGGCGAGAGGGAGCCAGAGGCTCTCTCTTTCGGG 1920
Db 1861 agtctgggaatgtaagcgcgaggtggtggcgaggagccagagactctctcttgggg 1920
QY 1921 TTTGGGGGCTTGACCTTGGGTGCGCTTTCTGACAGACTTTTACAGCCCCCGGGGSCACAGT 1980
Db 1921 ttttgggggcttgacctgggtgcgtcttctgacagactttacagccccggggcacagt 1980
QY 1981 CGTAGAGAGGGGGCGGGCGGCCATTTGGGGCTCTCTATTTGGGGTGTGTGGGGCGACCCC 2040
Db 1981 cgtagagaaggcgggcgccattgggctctctcattgggtgcttggggcgacccc 2040
QY 2041 ATCGGCTTACCGGCGCTCCCGAAATTTGGGGGAGCAAAAGGCTCTGCAGTCTCGGCTGAG 2100
Db 2041 atcggttaccggcggttccccggaattgtgggggacaaaagctctgcag tctcggtcgag 2100
QY 2101 GGTCTCACCGACAAAAGAGGGGAAGCCGGTGAGCAGAGGCTGAAAGGGTGGGGAAGCA 2160
Db 2101 ggtctccaccgacaaaagaggggaagccggtgagcagaggtggaagaggttggggaagca 2160
QY 2161 GGGGAGCTGTGCTGTCTCGGGCGGGTGGAACTTTAGCGGACCCCTGGGAGGAGGCTCCC 2220
Db 2161 ggggagctgtgcgtgtgtcgggcggggtggaaccttagcggaaccttgggagggaggtccc 2220
QY 2221 CGGCCAAACCTGCCGACCTCTCCCGCGGTGTGCTTTCAGAGCGCGCCACCGCC 2280
Db 2221 cggccgaacctgcagacctctctccccggcttgccttgcagggcgcgccccacgcgc 2280
QY 2281 CGCCGCTATGGGCTGAAGCCGCCAGAAAGACGCTGTTCGCCCTGCGCTCCATTCGACGA 2340
Db 2281 cgcgcctatgggctgaagccgccagaaagacgctgttccgcgtgcgtctccatcgagca 2340
QY 2341 CGTGTGCGGCTGTTCCTCCAGCTGGCGCGAGAGGAGCCGGACCTGCTGCTCTTTC 2400
Db 2341 cgtggtgcgctgttgcgtcgagctggcgagagagagagagagagagagagagagagag 2400
QY 2401 CTGTGTGCTGGGCTGTGGAGCATTTTCTGCTGTCAACCGGCTCATCTCCATCAACAGT 2460
Db 2401 ctttgggtgcgggttgcgtggagcatttcttggctgttcaacccgcgtctctccatccacag 2460
QY 2461 TCCGAGCTCACCTTCCAGCCAGCCCGCCCGCCGACCCCGCTGGCGGCTTCACCTACTT 2520
Db 2461 tcccgagctcaacttccagccagcccccgccccagacccgctggcgccctccacctact 2520
QY 2521 TCCGCTGGCGAGCTGTCTATCATCATGCGCGCCCTCTATGCGCGCTTACCGCCAGATCCG 2580
Db 2521 tcccgtggcgacgtgtctatcatcgccctctatgcccgcgttccagcgtccagagatcccg 2580
QY 2581 AGCGCGCTGACCTGTCTCCCTCTATCTCGAGAAGGGGTGTCTCCAGCGGTGAGCTGGT 2640
Db 2581 agcgcgctgcagcgtgtccctctctctcctcgagaaggggggtctctccagcgtctgagctg 2640
QY 2641 GAAGAAGGTCTCGGATGTCTATGGAACAGCCCTCAGCGCTCTCTACTTCAAGATCGGGC 2700
Db 2641 gaagaaggtctccgatgtcata tggaaacgctcagcgctctacttctcaagagatcgggc 2700
QY 2701 CCACATCCAGTCCCTCTTCAAGTTCATCAGGTTGGAGCCAGTAGGTGGGAATCTTAT 2760
Db 2701 cccatctcagtcctctctcagcttca tcaaggttggagccccagtaggtgggaatcttat 2760
QY 2761 CCATGACCCACTTCTTCAAAACCTCCATGTTTACAGAACCCTTTTAAAGACTGTAAAGC 2820
Db 2761 ccatgacccacttcttcaaaacctctccttgggttctacagaacctcttcaagaacttgttagc 2820
QY 2821 CTTTCTGAGGTTCGCGAGGTATTATTTTCTCTTTTGGAGTGGGAAACTGAAGCCACAGA 2880
Db 2821 ctttgtgaaggttcggcaggtgtattttctctcttctgagtttggaaactgaagccccagaga 2880
QY 2881 GGGGAATGATATGCCAAACTCACACGCGCATGGCAGGCTGGAACTGAAGCTGATCA 2940
Db 2881 ggggaaatgatafgccaaagtccacacgcatggcgaggtctggaaagtgaagcctgatca 2940

QY	2941	CTTGGCTCCAAATCATCAACCTCACTCTGCCCCCTCAGCACCCCCACCTTGGCCACTGA	3000
Db	2941		
		cttggctccaaatcatcaacctcaactgccccctcagcaccgccaccttgcacctga	3000
QY	3001	ACAGCTACAGGAGTTCTAAGCATGAGACACAGAGGGGGCAGCAGATTAGGGGGCAAGA	3060
Db	3001		
		acagctacaggagttcttaagcatgagacacagagagggcgagcaagattttaggggggcaaga	3060
QY	3061	AGATGAAATGGGGTGCAPTTGGAGCAGTTAAACAAAATAATGCTATGAAGATTTTTTT	3120
Db	3061		
		agatgaaattggcttgcatttgaggcagttaaacaaataatggctatgagattttttt	3120
QY	3121	TTTTTTTTTTTTTTTGAGACAGGGTCTCACTGTGTCGCCCGAGCTGAGTGCAGTGTGTG	3180
Db	3121		
		tttttttttttttgagacagggtctcactctgtcccccaggctggagtgcatgggtgg	3180
QY	3181	ATCATGGCTCACTGAGCCCTCAGTCTCGCTGGGCTCAGAGATCTCCAACCTCAGCCTCC	3240
Db	3181		
		atcatggctcactgagcctcagttctcctbaggctcagagatcctccaacctcagcctcc	3240
QY	3241	TGAGTAGCTGAGAGTACAGGCAATGCAACGCTGGTGGTGAATTTTTTTGATTTTTTTTG	3300
Db	3241		
		tgagtagctgaggtacaggcatgcacgtggctgggttaattttttgtattttttttt	3300
QY	3301	TACAGATGTGTCTCACTATGTGGCCAGACATGGTCTTGAACTCTTGGGCTCAAGTGAATC	3360
Db	3301		
		tagagatgggtctcactatgtggcccagagactggtctctgaaactcttgggtctcaagtgtatc	3360
QY	3361	TGCCCCCTCAGTCTCCCAAAATGCTGGGATTACAGGTGTGAGCCACCGCAACTGGTGGCC	3420
Db	3361		
		tgcgcgcctcagctctcccataatgtcgggatlacaggtgtgagccacgcgaactgggtggcc	3420
QY	3421	TATGAAAAATTTTTTTTTTTTACACGCGGCTCTACATCTCTGCCCCAGAGCTGGAGTGCA	3480
Db	3421		
		ta tgaataatttttttttttttcagacggcgctcaactctgtgcccaaggctggagtgca	3480
QY	3481	GTGGTGCAATTCGGGCTCACTCAAGCTCTGCCTCTGCTTTTCATGCCATTCCTCTGCCCT	3540
Db	3481		
		gtggtgcaattctggctcactgcaagctctgctcctgtcttcttgcaattctctctgct	3540
QY	3541	CTGTGCTCAGGCTCTGAGTAGCTGGGACTACAGAGCGCTGCCACCAATGCTCGCTAAAT	3600
Db	3541		
		cttgctcagcctctcctcagctagctgggtactacagagagcctggccaccatgctcggcctaatt	3600
QY	3601	TTTTTTTTTGGAATTTTAGTAGACAGCAGGTTTACCATGTTAGCCAGAGATGGTCTCGATCT	3660
Db	3601		
		tttttttggaatttttagtagacagaggttttcaacctgtagccagagatgggtctcgaatct	3660
QY	3661	CTTGACCTCGTGATCGCCCGGCTTGGCGCTCCCAAGTGCTGGGATTTACAGGCGTGAGCC	3720
Db	3661		
		cctgaectcgtgatccgcgcgccttggctcccgaagtgtcgtgggatlacagggogtgagcc	3720
QY	3721	ACCGCACTGGTCAAAAATTTTGTAGACAGAGAGGGGCTTGACCTCAAAAGGCTTAAGA	3780
Db	3721		
		accgcaacctgg tcaaaaaatgttttgagacagagagaggggcttgaaactcaaaagccttaaga	3780
QY	3781	GTACAGGCTTGCAAGAGCTTTTGCAACCAAGCCGGTTGACTTGGCAATCCCAATCCTTGGTGT	3840
Db	3781		
		gtcagggcttgaaagagcttttgaccaagccgggtgacttggcgaatccccatcctcgggtgt	3840
QY	3841	GCCATATTGAGAGGAATCAGAGGCTGTCTCTACGTTTAGCAGCAAAAGAGTGCAGAGAT	3900
Db	3841		
		gccattattgagaagggaatcagaggctgtctctcagcttagcaggaaagagtgagagat	3900
QY	3901	AAATGAGGGTTATTTCTGGTGGGTGTATAGCCAGAGAGTGTGGCCAGAGCTCTGTTTTT	3960
Db	3901		
		aaatgagggttatattgttgtgggtgtatagccagagagttgtggccagagctcctgttttt	3960
QY	3961	TGCCATTCCTGTTTTTAACCTAGTAGTGAGTAAATGGAATCCCTTAATCCATAGAATA	4020
Db	3961		
		tgcattctcgttttaactcagtagtaagtcagtaaagtgaatacctccataatccatagagata	4020

Qy	4021	TATAATAGACTTCCAGAGAAACAGCAGGTAGGGCCAAAGGCTTGUSTCAGCTACAGGAATAT	4080
Db	4021	tataatagagttgcagagaaacagcaggtagggccaaaggctgggctcagctacaggatct	
Qy	4081	CCAGAAAGGTATCTTGTGGACATAGAGGTTAAACAGGAGAGAGTCTTTGAACACGT	4140
Db	4081	ccagaaaggtatcttgtgacatagaggggtataaacaggagagagctcttgcacaacgt	
Qy	4141	GGAGGGAAGGGATGAGGGATAGTGGGCAGAGGAATCTGAGGTTGGTTCACAGGCTTGG	4200
Db	4141	ggagggaaaggatggagggatagtgggcaggagaaactgaggctgggtcacagagctgg	4200
Qy	4201	AAAGGGAGTGGGAGGAGTGTGGCCCATCAGTACCTGCTGCCCTTTGCCCATGTTAAAGCA	4260
Db	4201	aaagggaagtgggagggagtggtggcccatcactacctcctggcccttccccclgtlaaagca	4260
Qy	4261	CAGAGGACCTCTTTCAATCACTCCCTCTCCACAGGCACCAAAATGGACAGCTCCGGTG	4320
Db	4261	cagaggacctcttcatctacctccccctccacaggcaccaaatggacaagctccggtg	4320
Qy	4321	TGGCCTTGCTGTGGTGGGGCTGCCAGGCCCTGGGTCTCCGGGATGTCACACCTCGCCC	4380
Db	4321	tggccttgctgtggtggggcctccaggccctgggtctccggagtgccacctcgccc	4380
Qy	4381	TGTCAGGATCATGCTGGGTAGTGTTTGGGCCCAATGGGGAGCAGACAGCTGAGGTCA	4440
Db	4381	tgctcagatcatgctgggtagtggttggggcccaatggggagcagagctgaagtca	4440
Qy	4441	CTTGGCAGGCAAGGCAAGAGGACCGCAGGGGCCACAGTCAATGCCGTGGCTGTGGCTG	4500
Db	4441	cctggcaggcaaggcaagagacggcaggggccagagctcaatggcgggtgtggtctg	4500
Qy	4501	AGGGGTATTGTTCCCTCCCTCCAGCTGTGCCCTTCATCTACTGTAGTAGCCCAAGCCAC	4560
Db	4501	aggggtattgttccctccccagccttgttcccttcatactgtagtagcccaagccac	4560
Qy	4561	CCAAGGACTTCCATTTCTTGGGCCACACCCCTTTTCCCATCACACCACATAGGAA	4620
Db	4561	ccaaggacttccatttcttgggccacaccccttcttcccatcacaccacacataggaa	4620
Qy	4621	GGGAAGACACAGAGCCCTTTTCTTGGCTGTCAATTCCTGAAGCAGGCACAGGTTGGGC	4680
Db	4621	gggaagacagagagcccttctctgggtgtcaatctccctgaagcaggcacaggggtgggc	4680
Qy	4681	CATCATGAGACATAATGATCTCATCCCCCTTAAGAGTGGCTGTACCTGAAGAGTATAT	4740
Db	4681	catcatgagacataatgatctcatcccccttaagagctgggtgtacctgaaaggtacat	4740
Qy	4741	ACATGCCCTGTGACCCGACAGATGGAGTGGCGCTCATGGTGTGCCATCAACCTTCCA	4800
Db	4741	acatggcctgtgaccgcaagaagggtggcgctcatgggtgtgccatcaaccttccca	4800
Qy	4801	TTGACCTGTACACCGACTCGCTGGAGCTTCTGCAGCTGCAGCAGGTGAGGGCTGAGCCAA	4860
Db	4801	ttagcctgtcacacgagactcgcgtggagctcttcagctgcagcaggtgagggctgagccaa	4860
Qy	4861	TGGGGCAGGACTGGGCTAGCCAGACTTGACTTGCTGTGGAGCCCTGGGCAGGGGCACATT	4920
Db	4861	tggggcaggactgggctagccagacttgacttgctgtgggacccctgggagggggcactt	4920
Qy	4921	TCCCTCTCAGCTTCAGCTTCCCTCTCGGAAAATGGTTAGTAATKCTGGCCCTGGC	4980
Db	4921	tccctctcagcttcagcttcccccttctggaaaaatgggttagcaatctctggccttggc	4980
Qy	4981	CTTTCCCAGGGCTCTTTGGGAGAGTAGAATTGAGATGTGAAATTTGCTTTGACTTCCATTTAAA	5040
Db	4981	ctttcccaggctcttgggagagtagaaattgagatgtgaaattgcttctgactcattaaa	5040
Qy	5041	GGGCTGTCCCAAGAAATTTGGCCCTTCCACATGGTGGGTGCTCCTGTGTGGTCTTGACCC	5100
Db	5041	gggctgtcccaagaaatttggcccttccacatgggtgggtgctcctgtgtgtcttgaccc	5100
Qy	5101	CCACCTCTGCCGATAGGCTTAAGGACCCGCTTCTCCTCCTGTTCGCTGGCTCATTAAGTCT	5160

/*tag= b

FT XX PN W09839439-A1.
XX PD 11-SEP-1998.
XX PF 04-MAR-1998; 98WO-US04258.
XX PR 05-MAR-1997; 97US-0040269.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
PI Debelenko IV, Emmert-Buck MR, Guru SC, Liotta LA;
PI Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
XX WPI: 1998-506360/43.
DR P-PSDB; AAW29749.
XX
PT Protein and gene associated with multiple endocrine neoplasia type 1
PT - useful in gene therapy and to diagnose sufferers of, and those
PT susceptible to, this condition by detecting protein absence or gene
PT mutation(s)
XX
PS Claim 7; Page 55-56; 75pp; English.
XX
CC This is the transcribed sequence (cDNA) of human MEN1, a novel
CC tumour suppressor gene which is associated with multiple endocrine
CC neoplasia type 1. MEN1 cDNA can be obtained e.g. from leukocyte
CC RNA using MEN1-specific probes and PCR primers. A genomic sequence
CC (see AAW57416) is also provided. MEN1 encodes a 67.5 kDa protein
CC (see AAW28749), termed menin, that exhibits no apparent similarities
CC to previously known proteins. The lack of a functional menin
CC polypeptide, either by absence of the protein, its alteration
CC and/or associated mutations in the MEN1 gene, have been identified
CC in individuals with familial multiple endocrine neoplasia type 1
CC (FMEN1) and suffering from multiple endocrine neoplasia type 1.
CC The identification of MEN1 provides a new window into the mechanism
CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
CC MEN1 associated cancers, and provides preclinical identification of
CC individuals with the FMEN1 syndrome, i.e. cancer-free individuals
CC that are at high risk of acquiring MEN1 associated tumours. It
CC also provides a basis for gene therapy. The MEN1 nucleic acids may
CC also be used therapeutically to produce menin recombinantly and to
CC produce transgenic animals useful in research.
XX
SQ Sequence 2772 BP; 577 A; 909 C; 779 G; 507 T; 0 other;

Query Match 14.3%; Score 1312.2; DB 19; Length 2772;
Best Local Similarity 96.5%; Pred. No. 1.6e-234;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7499 GGATGCCAGACGAGGTCTGGAGTTCAGCCACTGGCGGCAACCTTGCTCTCACTT 7558
DB 1382 ggaaggcagtcacacgcctgtgtgcagtggtggcgaccccttctgtgagtcct 1441
QY 7559 GCTCTCCCACTGGCCAGGTGCGGAGAGGTGCGCATAGTGAGCGGAGCGCGAGGC 7618
DB 1442 aggcctttttagggagaggtgcgcagaaaggtcgcatagtgagccgagggccgaggc 1501
QY 7619 GGCCGAGGCGGAGGAGCGGTGGGGGAGGAGCCCGGAGGCGCGGCGGGCCACG 7678
DB 1502 ggcggagccgagagccgtgggcgaggaagcccggaagccgagggcccgagc 1561
QY 7679 GCGGGAGTCCAGCCAGAGAGGCCCCCGCCGCCAGAGACCCAGACTGGACAGGGCCT 7738
DB 1562 gcgggagtcgaagccagagagcccccgcgcgcccgaagaagccagcactggaagggcct 1621
QY 7739 GGCACCGCCGAGGAGGTGAGTGCAGGACCCCGGAGAGCCCTCCCTGGGAGCTGCTGG 7798
DB 1622 gggcaccggccaggggtgcagtggtccagagccccccccggagccctctggactgtcgtgg 1681

QY 7799 CACAGCCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 7858
DB 1682 cacagcccgagggcctggaaggtggcagcagcgtcaggtgcccagcaccgcagcacaacc 1741
QY 7859 ACCGCGGAGGGTCCAGTGTCTACTTTCCAGAGTGAGAGATGAAGGCGATGAAGAGCT 7918
DB 1742 accgcccggagggtccagtgctcacttccagagtgagaaagatgaaggagcatgaaggagct 1801
QY 7919 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAGCTCAGCTCAACTCAGCGCACAGTCGCA 7978
DB 1802 gctgggtggccaccaagatcaactcgagcgccatcaagctgcaactcacggcagcagtcgca 1861
QY 7979 AGTCAGATGAAGAAGCAGAAAGTGCCACCCTAGTACTAGTACTACACTCTGTCTTTCTCAA 8038
DB 1862 agtcagatgaagaagcagaaagtgtccacccttagtgactacactctgtcttctctcaa 1921
QY 8039 GCGGACAGCCAAAGGCTCTGAAGTACTGGGACTTCGGACCGCTTGTGGGGACCCAGGC 8098
DB 1922 gcggcagcgcaaggcctctgaactactgggacttcggaccgcttgtggggacccaggc 1981
QY 8099 TCGGCCTTAGTCCGCCCAACTCTCAGCCCAATGTCGCCCCAGCCCAAGGGGACAGGCC 8158
DB 1982 tcgcctctagtcgcccaactctgagcccatgttctgtccccagcccaaggggacagggcc 2041
QY 8159 TCACCTCTACCCAAACCTAGGTTCGCGGTCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
DB 2042 tcacctctacccaaaccttaggttcccggtcccgagtagactgtgtatcaaacccagat 2101
QY 8219 TTTCTCAGCTCAGAACCCAGGCGCTGTGCCCCAGTGGTGTAGATATAGGTCTCTTCTCC 8278
DB 2102 ttctctcagctcagaaacccagggctctgtgcccgagtcgttagaataaggtctctctccc 2161
QY 8279 AGAATCCAGCGCGCCCAATGGAACTCAGCTGGGTGGTCTTAATTACAGCTTTTAAAGGC 8338
DB 2162 agaatcccgccggcccaatggaaacctcagcgtgggtcttaattaccagtccttaaggcc 2221
QY 8339 CCAGCCCTTAGAAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 8398
DB 2222 ccagcccttagaaacccagctctctcgtgaaccgctcactcactagagcagacacagtt 2281
QY 8399 ACTCAGGCTCTCCAGCTGTGTAGAGCTGAGGTTCACCCCTTAACCCAAAGGAGGACACA 8458
DB 2282 actcagggctctcccagctgttagagctgaggtgttccaccttaacccagggagcaca 2341
QY 8459 GGTCCACCTCCAGCGCGGAGCCCTAGGACCACTCAGGCCCTTAGGAGTATATTTCGCA 8518
DB 2342 ggtccacactccagccggggagccttaggaccactcagcccttaggagtgatatttccgca 2401
QY 8519 CTTAGAAATCCATATCTTGCGAATCCAGCTCCCTGCCCAATAACTTCAGTCTCTGCT 8578
DB 2402 cttcagaattccatatctgtcgaatccaagctccctgcccacaataacttcagtcctgct 2461
QY 8579 TCCAGAAATTTGAAATCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
DB 2462 tcagaatttggaattcctagtttctctctctctctctctctctctctctctctctctctc 2521
QY 8639 CCGCCCCAGCTCTATGAGCATCTGAGCCCGCCGCTCTCTCTGACGAACCTGCGCCCGGA 8698
DB 2522 ccgccccagcctatgagcactcctgagccccctctctctctctctctctctctctctc 2581
QY 8699 TCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTTCCAGAGGTCCAGCCCATCTCGG 8758
DB 2582 tcagagcaggaaacctctctccgacctctggaaacctccccagaggctccagcccatctcgg 2641
QY 8759 AGCATCCCGAGGAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAGCTGATCTCTTC 8818
DB 2642 agcatccccgaggaactctgcagaggggttaggagtggtgacaagagcctgatctcttc 2701
QY 8819 CTCTTTGTACATAGATTTATTTTTCAGTTCCCAAGAAAGATGAATACATTTTGTAAAAA 8878
DB 2702 ctgtttgtacatagattttatttttcagttccccagaagaagatgaatacatatttgttataaaa 2761
QY 8879 AAATATAAA 8887

RESULT 7
AAV27941/c
ID AAV27941 standard; cDNA; 14796 BP.
XX AC AAV27941;
XX DT 25-SEP-1998 (first entry)
XX DE Survivin gene.
XX KW survivin; apoptosis; cellular apoptosis; transplantation; ss;
KW motor neuron degenerative disease; HIV infection; immunosuppression;
KW gastrointestinal perturbations; cardiovascular disorder.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 2811..12044
FT FT /*tag= a
FT FT /product= Survivin
FT FT 2811..2921
FT FT /*tag= b
FT FT /number= 1
FT FT 2922..3173
FT FT /*tag= c
FT FT /number= 1
FT FT 3174..3283
FT FT /*tag= d
FT FT /number= 2
FT FT 3284..5157
FT FT /*tag= e
FT FT /number= 2
FT FT 5158..5275
FT FT /*tag= f
FT FT /number= 3
FT FT 5276..11954
FT FT /*tag= g
FT FT /number= 3
FT FT 11955..12044
FT FT /*tag= h
FT FT /number= 4
PN WO9822589-A2.
XX 28-MAY-1998.
XX 20-NOV-1997; 97WO-US21880.
XX 20-NOV-1997; 97US-0975080.
PR 20-NOV-1996; 96US-0031435.
XX (UYVA) UNIV YALE.
XX Altieri DC;
XX WPI; 1998-312475/27.
DR P-PSDB; W614359.
XX Modulating apoptosis by controlling the Survivin gene - useful for
PT treating transplant rejection, degenerative disorders and tumours
XX Disclosure; Fig 10; 108pp; English.
XX The survivin gene can be used to control apoptosis through modification
CC of the gene. Survivin peptides can be used to inhibit cellular
CC apoptosis, e.g. for enhancing the viability of organs and tissues prior
CC to their transplantation, for preserving the growth of cells in culture
CC or for treating conditions involving abnormal apoptosis,
CC e.g. degenerative diseases such as motor neuron degenerative diseases,
CC HIV infection, dermatological effects of ageing, disorders and diseases
CC such as immunosuppression, gastrointestinal perturbations, cardiovascular
CC disorders, apoptosis related to reperfusion damage, rejection of tissue
CC transplantation and Alzheimer's disease. Agents which block Survivin

CC activity can be used to treat e.g. tumours.
XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 other.
SQ
Query Match 3.8%; Score 348; DB 19; Length 14796;
Best Local Similarity 66.8%; Pred. No. 1.8e-55;
Matches 624; Conservative 0; Mismatches 285; Indels 25; Gaps 8;
QY 382 GCACGACAGTGGCTCACACCTGTATCCACAGCCTTTTGGAGGCGGAGGTAGGAGGAT 441
DB 6289 GCGGGTGGTGGCTCACACCTGTATCCACAGCCTTTGGAGGCGGAGGAGGAGAT 6230
QY 442 CACTTGAGGTGAGGAGTTCAGACGCTTGACCATGGTGACACCCCTGTCTCTACT 501
DB 6229 CAC--GAGGTGAGGAGTCAAGACCTCTCTGGCTAACACGGTGAACACCG-GTCTCTACT 6173
QY 502 AAGTAAATACAAAATTAGCCAAAGTGTGTGGCAGCAGCTGTAAATCCGGGTACTTTGG 561
DB 6172 AA-AAATACAAAATTTAGCCGGGTGTGTGGGCGGCGCTGTAGTCCAGCTACTCGG 6114
QY 562 GAAGCTGAGGAGAGAAATCACTTGAACCTGGGAGGAGAGGTTCAGTGCAGCGAGATC 621
DB 6113 GAGGCTGAGGAGGAGAAATGGCTTGAACCTGGGCGGCGAGCTTGCAGTGCAGCGAGATC 6054
QY 622 ACCCCACTGCACCTCCAGCCTGAGTGACAGAGCGAGACTCTGTCTCAAAACAA----- 673
DB 6053 CCGTCGCTGCACCTCTAGCCTGGGCGACAGAGCAAGACTCCATCTCAATAAATAATAA 5994
QY 674 ---ATAACAATAACTACTCTTTGGCCGGTAAAGTGTGTTCACGCTGTAAATTTAGCA 730
DB 5993 TTTAAAAATAAAATACTTCCCTGAGGCCAGGTGTGTGGTGTACACCTATAATCTCAGCA 5734
QY 731 CTTTGGAGGCTGAGGCGGCGACATCACTTGAAGTTAGGGTTTCAGACACCACTTCGCCA 790
DB 5933 CTTTGGAGGCTGAGGCGGCGTGGATCACCCTGAGGTGAGGAGTTCAAAACACGCTGGCCA 5874
QY 791 ACATGTTGAAACCCCATCTCTACTTAAATAACAAAAGTTTCTGGGTGTGGTGGCGGAC 850
DB 5873 ACATGTTGAAACCTCGTCTCTACCAAAAATAACAAA--TTACCTGGTGTGGTGGCACAT 5815
QY 851 GCCTATAATCCAGCTACTTTGGGACTTTTAAAGACGGAATCTCTACTCTGTGGCCAG 910
DB 5814 GCCTGTAATCTCGGCTACTTTGGGAGGCT-----GAAGCAGGCAAAATCACTTGAGCCCA 5762
QY 911 GCTGGAGTGCAGTGGCAAGATTCTGGCTCACTGAAGCTCCGCTCCAGGTTCAAGGGG 970
DB 5761 GGAGCGGAGGTTTGCAGTGTATCCGAGATCAGGCCACCTACTCCAGCTGGTAAACAGAG 5702
QY 971 ATTCCCGCGGCTCAGCCTCCCAAGTAGCTGGGAATCCCTGTCTCTGCAAAAAAATAA 1030
DB 5701 TGAACCTGTCTCAGAACAAAAAAGCCAGCGGTGGTGGTTCAGCCCTGTAATC 5642
QY 1031 AAAAAAATAAATAAT 1090
DB 5641 CCAGCACTTTGGGAGCCCAAGGTGGCGGAATCAGAGGTTCAGGAGTTCAAGACCAAGCTG 5582
QY 1091 TTATATGTAT 1150
DB 5581 GGCAATGTGTGTAAGAACCCCATCTCTACTTAAAAATACA-AAATATAGCTGGTGGTGGC 5523
QY 1151 GCACGCTGTGTATCAGATCAGCTTACTCGGAGGCTGAGGAGGAGGAGGAGGAGGAGG 1210
DB 5522 ACTTCCTCTGTAATCCAGTTACTTTGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGG 5463
QY 1211 AGTCGAGGTGAGTGCAGGCGGAGTGCAGCTGAGTGCAGTTCAGCTGGGCGGA-AAGAGAA 1269
DB 5462 GGCAGAGGTTGAGTGCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5403
QY 1270 AGACCTGTCTCAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1303
DB 5402 AAACCTATCTCAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5369

[illegible]

OS	Homo sapiens.
XX	
PN	WO200116311-A1.
XX	
PD	08-MAR-2001.
XX	
PF	
XX	31-AUG-2000; 2000WO-JP05930.
XX	
PR	31-AUG-1999; 99JP-0245962.
PR	09-MAY-2000; 2000JP-0136266.
XX	
XX	(HISM) HISAMITSU PHARM CO LTD.
PA	(CHIB-) CHIBA PREFECTURE.
XX	
PI	Nakagawara A;
PI	
XX	
DR	WPI; 2001-226686/23.
XX	
XX	
PT	Human lp36 homozygosity deletion domain from the 36-position of first
PT	chromosome short arm in human neuroblastoma cell lines, applicable e.g.
PT	in gene diagnosis of tumors as well as in developing anti-cancer drugs
PT	
XX	
XX	
XX	Example 8; Page 67-88; 226pp; Japanese.
XX	
CC	The present invention describes a homozygosity deletion domain
CC	co-existing in the 36-position of the first chromosome short arm (lp36)
CC	in human neuroblastoma. Also described are base sequences from the lp36
CC	position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1),
CC	which are tumour suppressor genes in human neuroblastoma. The genes are
CC	tumour suppressor genes, base sequence data of which are applicable as
CC	tumour markers and reagents in studying mechanism of tumour body
CC	formation, and gene diagnosis of tumours as well as in developing
CC	anti-cancer drugs. AAF97787 to AAF97829 represent PCR primers used in
CC	the exemplification of the present invention, and AAF97830 to AAF97874
CC	represent sequences given in the exemplification of the present
CC	invention.
XX	
XX	Sequence 51474 BP; 14008 A; 11363 C; 11657 G; 14446 T; 0 other;
SQ	

Query Match	3.7%	Score	342.8;	DB	22;	Length	51474;
Best Local Similarity	78.8%;	Pred. No.	2.1e-54;				
Matches	490;	Conservative	0;	Mismatches	112;	Indels	20;
Gaps	6;						

Qy	3121	TTTTTTTTTTTTTTTGACAGGGTCTCAGCTGTGCCCCCAGGCTGGAGTGCAGTGGTGTG	3180
Db	15291	TTTTTTTTTTTTTTTGACAGAGTCTCGCTGTGCCCCAGGCTGGAGTGCAGTGGGCGG	15232
Qy	3181	ATCATGGGCTCACTCGAGCCTCAGTCTCCCTGGGCTCAGAGATCTCTCCAACTCAGCCTCC	3240
Db	15231	ATCTCGGCTCACTGCAAGCTCCGGCTCCAGGTTCAAGCCATTCCTCGCTCAGCCTCC	15172
Qy	3241	TGAGTAGCTGAGATCAGGGATCGACGGTGGTG-CTGGTTAAATTTTTTTGATTTTTTTT	3299
Db	15171	CGAGTAGCTGGGACTACAGGCACCTGCCACCATGCCGGCTTAATTTTTTTGTATTATTA	15112
Qy	3300	GTACAGATGGTGTCTCACTATGTGCCCAGACTGGTCTTTGAACCTCTGGGCTCAAGTCAT	3359
Db	15111	GTAGAGATGGGGTTTCACCATGTTAGCCAGGATGGTCTGCATCTCCTGACCTC--GTGAT	15054
Qy	3360	CTGCCCGCCTCAGTCTCCC-AAATGCTGGGATTACAGGTGTGAGCCACCGCAACTGGTGG	3418
Db	15053	CCGCCCGCCTCGACCTCCCAAATGGTGGGATTACAGGCATGAGCCACTGCGCCAGCCT	14994
Qy	3419	CCTATGAAAATTTTTTTTTTTTTCAGACGGGCTCTACTCTGTGCGCCAGGCTGGAGTG	3478
Db	14993	ATTTT-----TTTTTTTTTTTCTTGAGATGAGGCTCGCTGTGTTGCCAGGCTGGAGTG	14939
Qy	3479	CAGTGGTGCAATCTCGGCTCACTGCAAGCTCTGCCCTCTGCTTTTCATGCCATTCTCCTGC	3538
Db	14938	CAGTGGGCACAATCTCGGCTCACTGAAACCTCCGGCTCCAAAGTTCAAGCGATT-----	14884

Qy	3539	CTCTGCGCTCAGCCTCTGAGTAGTGGAGACTACAGGAGCGCTGCCACCATGCCCTGGCTAA	3598
Db	14885	CTCTGCGCTCAGCCTCTGAGTAGTGGAGACTACAGGAGCGCTGCCACCATGCCAGCTAA	14892
Qy	3599	TTTTTTTGGGATTTTAGTAGAGACGAGGTTTCACCATGTTAGCCAGGATGGTCTCGAT	3658
Db	14825	----CTTTTGTATTTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGATGGTCTTGAT	14770
Qy	3659	CTCCTGACCTCGTGATCGCCCGCGCTTGGCCTCCCAAAGTCTGGGATTACAGGCGTGAG	3718
Db	14769	CTTTTGACCTCGTGATTTGCCCCACCTTGGCGCTCCCAAAGTCTGGGATTACAGGCGTGAG	14710
Qy	3719	CCACGGCACCTGGTCAAAAATG	3740
Db	14709	CCATTGCACCGAGCTACAGTG	14688
RESULT 15			
AAV27941			
ID	AAV27941 standard; cDNA; 14796 BP.		
XX	XX		
AC	AAV27941;		
XX	XX		
DT	25-SEP-1998 (first entry)		
XX	XX		
DE	Survivin gene.		
XX	XX		
KW	survivin; apoptosis; cellular apoptosis; transplantation; ss;		
KW	motor neuron degenerative disease; HIV infection; immunosuppression;		
KW	gastrointestinal perturbations; cardiovascular disorder.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	2811..12044		
FT	/*tag= a		
ET	/product= Survivin		
FT	2811..2921		
FT	/*tag= b		
FT	/number= 1		
FT	2922..3173		
FT	/*tag= c		
FT	/number= 1		
FT	3174..3283		
FT	/*tag= d		
FT	/number= 2		
FT	3284..5157		
FT	/*tag= e		
FT	/number= 2		
FT	5158..5275		
FT	/*tag= f		
FT	/number= 3		
FT	5276..11954		
FT	/*tag= g		
FT	/number= 3		
FT	11955..12044		
FT	/*tag= h		
FT	/number= 4		
XX	XX		
PN	W09822589-A2.		
XX	XX		
PD	28-MAY-1998.		
XX	XX		
PF	20-NOV-1997; 97WO-US21880.		
XX	XX		
PR	20-NOV-1997; 97US-0975080.		
PR	20-NOV-1996; 96US-0031435.		
XX	XX		
PA	(UYVA) UNIV YALE.		
XX	XX		
PI	Altieri DC;		
XX	XX		
DR	WPI; 1998-312475/27.		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 01:32:59 ; Search time 15359.2 Seconds
(without alignments)
9244.854 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGGCTTGTAACCTCTGGCC.....AGCCACGGCGCGCGCCGCC 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htgo_hum1.*
23: em_htgo_hum2.*
24: em_htgo_hum3.*
25: em_htgo_hum4.*
26: em_htgo_hum5.*
27: em_htgo_hum6.*
28: em_htgo_hum7.*
29: em_htgo_hum8.*
30: em_htgo_inv1.*
31: em_htgo_inv2.*
32: em_htgo_other.*
33: em_htgo_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

- em_ov.*
44: em_pat.*
45: em_ph.*
46: em_pl.*
47: em_ro.*
48: em_sts.*
49: em_sy.*
50: em_un.*
51: em_vi.*
52: gb_sts1.*
53: gb_sts2.*
54: gb_sts3.*
55: gb_sy.*
56: gb_un.*
57: gb_vil.*
58: gb_vil.*
59: gb_vil.*
60: gb_vil.*
61: gb_vil.*
62: gb_vil.*
63: gb_vil.*
64: gb_vil.*
65: gb_vil.*
66: gb_vil.*
67: gb_vil.*
68: gb_vil.*
69: gb_vil.*
70: gb_vil.*
71: gb_vil.*
72: gb_vil.*
73: gb_vil.*
74: gb_vil.*
75: gb_vil.*
76: gb_vil.*
77: gb_vil.*
78: gb_vil.*
79: gb_vil.*
80: gb_vil.*
81: gb_vil.*
82: gb_vil.*
83: gb_vil.*
84: gb_vil.*
85: gb_vil.*
86: gb_vil.*
87: gb_vil.*
88: gb_vil.*
89: gb_vil.*
90: gb_vil.*
91: gb_vil.*
92: gb_vil.*
93: gb_vil.*
94: gb_vil.*
95: gb_vil.*
96: gb_vil.*
97: gb_vil.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	9179	100.0	9180	97	HSU93237
C 2	9042	98.5	203300	85	AC000134
3	8004.8	87.2	109290	84	HS838L14
C 4	5844	63.7	171980	82	AP000928
C 5	5072.2	55.3	73431	82	AP000575
C 6	4368.4	47.6	166269	82	AP001187
7	4118.2	44.9	157454	82	AP001462
8	3386.2	36.9	178100	82	AP001558

9	2071.4	22.6	73431	82	AP000575	Homo sapi
10	1982.2	21.6	166269	82	AP001187	Homo sapi
11	1312.2	14.3	2772	97	U93236	Human menin
12	1253.2	13.7	1691	91	BC002664	Homo sapi
13	1251.6	13.6	2633	91	BC002544	Homo sapi
14	1242.4	13.5	1837	97	HSURF	
15	1075.8	11.7	9286	94	AF109390	Homo sapi
16	1075.8	11.7	132365	94	AC006956	Mus muscu
17	1058.2	11.5	6779	94	AF093756	Mus muscu
18	1036.6	11.3	6736	94	AF024513	Mus muscu
19	942.4	10.3	171980	82	AP000928	Homo sapi
20	942.	10.3	178100	82	AP001558	Homo sapi
21	782.6	8.5	157454	82	AP001462	Homo sapi
22	782.6	8.5	938	93	HS297487	Homo sapi
23	749.4	8.2	2196	9	AR081879	Sequence
24	746.4	8.1	812	9	AR081882	Sequence
25	565.2	6.2	71297	78	AC091082	Homo sapi
26	545.8	5.9	71297	78	AC091082	Homo sapi
27	543.6	5.9	52358	77	AC090316	Homo sapi
28	536.6	5.8	3067	94	AB023400	Rattus no
29	525	5.7	2984	94	AF109389	Mus muscu
30	513.4	5.6	2947	94	AF130368	Mus muscu
31	483.4	5.3	944	93	HS297488	Homo sapi
32	477.8	5.2	2261	94	AF130369	Rattus no
33	471.4	5.1	538	9	AR081880	Sequence
34	466	5.1	501	93	HS297486	Homo sapi
35	465.6	5.1	816	93	HS297489	Homo sapi
36	465.4	5.1	513	93	HS297485	Homo sapi
37	456.2	5.0	2623	94	AF016398	Mus muscu
38	428.8	4.7	2577	94	AF072755	Mus muscu
39	408.8	4.5	1964	94	AB023401	Rattus no
40	396.6	4.3	1906	94	AF130370	Mus muscu
41	395.8	4.3	181683	66	AC021618	Homo sapi
42	395.4	4.3	49868	78	AC090368	Homo sapi
43	393.4	4.3	207091	77	AC087289	Homo sapi
44	393	4.3	154978	75	AC073952	Homo sapi
45	393	4.3	200491	86	AC007249	Homo sapi
ALIGNMENTS						
RESULT	1	HSU93237	9180 bp	DNA	PRI	19-APR-1997
LOCUS		Human menin (MEN1) gene, complete cds.				
DEFINITION		U93237				
ACCESSION		U93237.1	GI:1945388			
VERSION						
KEYWORDS		human.				
SOURCE		human.				
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 9180)				
AUTHORS		Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.-E., Collins,F.S., Emmert-Buck,M.R., Debeelenko,L.V., Zhuang,Z., Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A., Weisenmann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.				
TITLE		Positional cloning of the gene for multiple endocrine neoplasia-type 1				
JOURNAL		Science 276 (5311), 404-407 (1997)				
MEDLINE		97258940				
REFERENCE		2 (bases 1 to 9180)				
AUTHORS		Collins,F.S.				
TITLE		Direct Submission				
JOURNAL		Submitted (13-MAR-1997) National Human Genome Research Institute, Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD 20892, USA				
FEATURES		Location/Qualifiers				
source		1. .9180				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
repeat_region						
gene						
repeat_region						
repeat_region						
repeat_region						
exon						
mRNA						
mRNA						
exon						
exon						
CDS						
variation						

```
/standard_name="L22R"  
/note="Lys22Arg"  
/replace="g"  
2534..2537  
/gene="MEN1"  
/standard_name="357del4"  
/note="4 bp deletion"  
/replace=""  
2593  
/gene="MEN1"  
/standard_name="416delC"  
/note="1 bp deletion"  
/replace=""  
2641..2643  
/gene="MEN1"  
/standard_name="K119del"  
/note="3 bp deletion"  
/replace=""  
2689  
/gene="MEN1"  
/standard_name="512delC"  
/note="1 bp deletion"  
/replace=""  
complement(3117..3725)  
/rpt_type="dispersed"  
/rpt_family="ALU"  
4297..4505  
/gene="MEN1"  
/number=3  
4445  
/gene="MEN1"  
/standard_name="W198X"  
/note="Trp198stop"  
/replace="a"  
4476..4479  
/gene="MEN1"  
/standard_name="735del4"  
/note="4 bp deletion"  
/replace=""  
4716..4844  
/gene="MEN1"  
/number=4  
5177..5217  
/gene="MEN1"  
/number=5  
5298..5385  
/gene="MEN1"  
/number=6  
6025..6161  
/gene="MEN1"  
/number=7  
6134  
/gene="MEN1"  
/standard_name="1132delG"  
/note="1 bp deletion"  
/replace=""  
6623..6758  
/gene="MEN1"  
/number=8  
6660..6662  
/gene="MEN1"  
/standard_name="E363del"  
/note="3 bp deletion"  
/replace=""  
7196..7360  
/gene="MEN1"  
/number=9  
7316  
/gene="MEN1"  
/standard_name="W436R"  
/note="Trp436Arg"  
/replace="c"  
7317
```

```
/gene="MEN1"  
/standard_name="W436X"  
/note="Trp436stop"  
/replace="a"  
7578..7874  
/gene="MEN1"  
/number=10  
7806  
/gene="MEN1"  
/standard_name="R527X"  
/note="Arg527stop"  
/replace="t"  
BASE COUNT 1936 a 2604 c 2612 g 2027 t 1 others  
ORIGIN  
  
Query Match 100.0%; Score 9179; DB 97; Length 9180;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 9180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CTGGTCTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAGTCTTGTAA 60  
Db 1 CTGGTCTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAGTCTTGTAA 60  
  
QY 61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATACATGCTCAGGTATTG 120  
Db 61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATACATGCTCAGGTATTG 120  
  
QY 121 TGAATGTCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGAGGAGATTATAA 180  
Db 121 TGAATGTCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGAGGAGATTATAA 180  
  
QY 181 ATTTTGGAGAAATCCAGAGAGGTGAAGAGTGAAGTAGGGCAGAAATTTAATCTGTTT 240  
Db 181 ATTTTGGAGAAATCCAGAGAGGTGAAGAGTGAAGTAGGGCAGAAATTTAATCTGTTT 240  
  
QY 241 ATTTACTGTATATACCGAGTGTCTGGAACCTTGCCCATGGTAAGTACCAAAATCTGTT 300  
Db 241 ATTTACTGTATATACCGAGTGTCTGGAACCTTGCCCATGGTAAGTACCAAAATCTGTT 300  
  
QY 301 TTTTGTGAATGAATAAGCAATAAATAGTGACCGTGGAAATTTAGTATTATTTCAAAGT 360  
Db 301 TTTTGTGAATGAATAAGCAATAAATAGTGACCGTGGAAATTTAGTATTATTTCAAAGT 360  
  
QY 361 TTCAAAGCGTTGTGATACAGGCCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTT 420  
Db 361 TTCAAAGCGTTGTGATACAGGCCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTT 420  
  
QY 421 TGGAGGCCGAGGTAGGAGGATCACTTGAAGTCAGGAGTTTCGAGACGAGCTGACCAACAT 480  
Db 421 TGGAGGCCGAGGTAGGAGGATCACTTGAAGTCAGGAGTTTCGAGACGAGCTGACCAACAT 480  
  
QY 481 GGTGACACCCCTGTCTCTACTAAGTAAAAATACAAAATTAGCCCAAGTGTGGTGGCAGGCA 540  
Db 481 GGTGACACCCCTGTCTCTACTAAGTAAAAATACAAAATTAGCCCAAGTGTGGTGGCAGGCA 540  
  
QY 541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAGAAATCACTTGAACCTGGGAGGCAG 600  
Db 541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAGAAATCACTTGAACCTGGGAGGCAG 600  
  
QY 601 AGGTTGAGTGAGCCGAGATCACCCACCTGACCTGAGCTGAGTGCAGACGAGCACTC 660  
Db 601 AGGTTGAGTGAGCCGAGATCACCCACCTGACCTGAGCTGAGTGCAGACGAGCACTC 660  
  
QY 661 TGTCTCAAAACAAATAAACAATAAATACTACTCTTTTGGCCGGTAAAGTGGTTCACGCCGTGT 720  
Db 661 TGTCTCAAAACAAATAAACAATAAATACTACTCTTTTGGCCGGTAAAGTGGTTCACGCCGTGT 720  
  
QY 721 AATTTTAGCACTTTGGGAGGCTGAGGGGCGAGATCACTTCAGGTTAGGGGTTTCGAGACC 780  
Db 721 AATTTTAGCACTTTGGGAGGCTGAGGGGCGAGATCACTTCAGGTTAGGGGTTTCGAGACC 780  
  
QY 781 AGTCTGGCCCAACATGTTGAAACCCCATCTCTACTTAAATACAAAAGTTTTCCTGGGTGT 840
```


|||||
Db 5161 CTCCTCGGCTCCTAGAAAGCTGCTGGCTGCTCTATGACCTGGAGACATCTGGAAGGTC 5220
|||||
Qy 5221 AGTAGAGGAAGTGGCCAGGCTGGCCCTGTGTAGGCGGGGGGCTGGGTGGCAGCCCTGAA 5280
|||||
Db 5221 AGTAGAGGAAGTGGCCAGGCTGGCCCTGTGTAGGCGGGGGGCTGGGTGGCAGCCCTGAA 5280
|||||
Qy 5281 TTATGATCCTTTTCTAGTACCCCATGGCCCTTAGGGAACCTGGCAGATCTAGAGGAGCTG 5340
|||||
Db 5281 TTATGATCCTTTTCTAGTACCCCATGGCCCTTAGGGAACCTGGCAGATCTAGAGGAGCTG 5340
|||||
Qy 5341 GAGCCACCCCTGGCCGGCCAGACCCAGCTACCCCTCTACCAAGGTGGGGGCACTCTAAG 5400
|||||
Db 5341 GAGCCACCCCTGGCCGGCCAGACCCAGCTACCCCTCTACCAAGGTGGGGGCACTCTAAG 5400
|||||
Qy 5401 GAGGCTGCAGAAAGGAGACCCCTAAAGTGGCTGAGGAGGGGCCCTCATCTGGGCAGATG 5460
|||||
Db 5401 GAGGCTGCAGAAAGGAGACCCCTAAAGTGGCTGAGGAGGGGCCCTCATCTGGGCAGATG 5460
|||||
Qy 5461 AGAAGAGAACTTTTGTGTGTTGGGGGATCGCCCATCCAGTCTCACTTTGTGTCAACTGT 5520
|||||
Db 5461 AGAAGAGAACTTTTGTGTGTTGGGGGATCGCCCATCCAGTCTCACTTTGTGTCAACTGT 5520
|||||
Qy 5521 GTGCAGATCAGTTTCACTCAGGCTGTCTGAGGGGTGTCCAGGGTTCCCGACGCTGGGAG 5580
|||||
Db 5521 GTGCAGATCAGTTTCACTCAGGCTGTCTGAGGGGTGTCCAGGGTTCCCGACGCTGGGAG 5580
|||||
Qy 5581 TGGCAGGGGCTGCAATTTGTCGCCCTCAGCCCTGCGCTTTTCTGCCACTGCTTACTGTCTTC 5640
|||||
Db 5581 TGGCAGGGGCTGCAATTTGTCGCCCTCAGCCCTGCGCTTTTCTGCCACTGCTTACTGTCTTC 5640
|||||
Qy 5641 CTGGAGTATAACAGAGGTCAAAATGTGTGTAGGACACTGAAGAGGGGTGTTCACCTTGGTG 5700
|||||
Db 5641 CTGGAGTATAACAGAGGTCAAAATGTGTGTAGGACACTGAAGAGGGGTGTTCACCTTGGTG 5700
|||||
Qy 5701 GGTGTAGGTGGGAGGAGGCGCATTTGGCTGGGCTTGAAAGTCTTTGGTGATGTGTAGAA 5760
|||||
Db 5701 GGTGTAGGTGGGAGGAGGCGCATTTGGCTGGGCTTGAAAGTCTTTGGTGATGTGTAGAA 5760
|||||
Qy 5761 CAGTGTCTTGAGAAAGAGAAAGGGCCCTGAGCTCGGAGGGCGAGGCCACCCCTGCAGTCTG 5820
|||||
Db 5761 GAGTGTCTTGAGAAAGAGAAAGGGCCCTGAGCTCGGAGGGCGAGGCCACCCCTGCAGTCTG 5820
|||||
Qy 5821 CCCAGGGCTCAGCCAGCAGTCTGTAGACCCAGGAGGAGACAGGTAGAAGGGCTGGC 5880
|||||
Db 5821 CCCAGGGCTCAGCCAGCAGTCTGTAGACCCAGGAGGAGACAGGTAGAAGGGCTGGC 5880
|||||
Qy 5881 AGCGAGTGGAGGTGGAGTGGAGATGAGAGGACTCCCTGGGATCTTCTGTGGCCCTT 5940
|||||
Db 5881 AGCGAGTGGAGGTGGAGTGGAGATGAGAGGACTCCCTGGGATCTTCTGTGGCCCTT 5940
|||||
Qy 5941 CTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGGTGCCCTCCCTGAGG 6000
|||||
Db 5941 CTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGGTGCCCTCCCTGAGG 6000
|||||
Qy 6001 ATCTCTGCTCCTACCTCCATCCAGGGCATATGCCCTAGCCAGACCTACTATPCGGGATGAA 6060
|||||
Db 6001 ATCTCTGCTCCTACCTCCATCCAGGGCATATGCCCTAGCCAGACCTACTATPCGGGATGAA 6060
|||||
Qy 6061 CACATCTACCCCTACATGTACTGGCTGGCTTACCACCTGTCCACACCGCAATGTGCGGGAA 6120
|||||
Db 6061 CACATCTACCCCTACATGTACTGGCTGGCTTACCACCTGTCCACACCGCAATGTGCGGGAA 6120
|||||
Qy 6121 GCCCTGAGGGCTGGGGGACACGGCCACTGTCTATCCAGGAGTGAGGATCCCCCTACTAG 6180
|||||
Db 6121 GCCCTGAGGGCTGGGGGACACGGCCACTGTCTATCCAGGAGTGAGGATCCCCCTACTAG 6180
|||||
Qy 6181 GGCTGACGCTGTCTTTTCTTCCCTCCATCAGTTTCCAAACACCCCTGTGTCCAGGACTG 6240
|||||
Db 6181 GGCTGACGCTGTCTTTTCTTCCCTCCATCAGTTTCCAAACACCCCTGTGTCCAGGACTG 6240
|||||
Qy 6241 AGGCTGGCTCCAGCCCATCCCTTCCATCCAGTCCCTAGGACGACGACCAT 6300
|||||

Db 6241 AGGCTGGCTCCAGCCGCCCATCCCTTTCCATCCAGTCCCTAGGACGACGACCAT 6300
|||||
Qy 6301 TACCAGAGGTAGGAGCCCTGATTAAAGGTGTACATCTTTCCCTCCCTCCCTCTCCTC 6360
|||||
Db 6301 TACCAGAGGTAGGAGCCCTGATTAAAGGTGTACATCTTTCCCTCCCTCCCTCTCCTC 6360
|||||
Qy 6361 CTAATTTTTTTTTTCTCAGAAACAGTCTCAAAATCTCCAATGTTAAACACCATCATCCAGC 6420
|||||
Db 6361 CTAATTTTTTTTTTCTCAGAAACAGTCTCAAAATCTCCAATGTTAAACACCATCATCCAGC 6420
|||||
Qy 6421 AGTGGGACTTCCACCCCTGGGCCCATGCCCCCTCTCTCATTTCTTCTTCTTCTTGGG 6480
|||||
Db 6421 AGTGGGACTTCCACCCCTGGGCCCATGCCCCCTCTCTCATTTCTTCTTCTTCTTGGG 6480
|||||
Qy 6481 CTGACCCAGACAGCATCATTTTGCAGTGTAGGACCCCACTACTCCCCAGGCCCTGGGG 6540
|||||
Db 6481 CTGACCCAGACAGCATCATTTTGCAGTGTAGGACCCCACTACTCCCCAGGCCCTGGGG 6540
|||||
Qy 6541 CTCATCCCCCGCCAGGTCTCTGGGGCTACCCCCGATGGTGAGACCCCTTCAGACCCCTAC 6600
|||||
Db 6541 CTCATCCCCCGCCAGGTCTCTGGGGCTACCCCCGATGGTGAGACCCCTTCAGACCCCTAC 6600
|||||
Qy 6601 AGACACCCCTGCTCTCACAGCTACACTACTGCCGGAAGACGAGCATCTACAAGG 6660
|||||
Db 6601 AGACACCCCTGCTCTCACAGCTACACTACTGCCGGAAGACGAGCATCTACAAGG 6660
|||||
Qy 6661 AGTTCTTTGAAGTAGCCAAATGATGTATCCCAACCTGCTGAAGGAGGACGACGCTTGC 6720
|||||
Db 6661 AGTTCTTTGAAGTAGCCAAATGATGTATCCCAACCTGCTGAAGGAGGACGACGCTTGC 6720
|||||
Qy 6721 TGGAGGGGGGAGGAGGGCCGGGGGAGCAAAAGCAGGTGAAAGGTGGAGCTTCCAGCC 6780
|||||
Db 6721 TGGAGGGGGGAGGAGGGCCGGGGGAGCAAAAGCAGGTGAAAGGTGGAGCTTCCAGCC 6780
|||||
Qy 6781 TGTCTCAGCTCCCACTGGACAGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA 6840
|||||
Db 6781 TGTCTCAGCTCCCACTGGACAGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA 6840
|||||
Qy 6841 CGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTGTCT 6900
|||||
Db 6841 CGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTGTCT 6900
|||||
Qy 6901 ATAAGTGTAGGTCTGGGGCCAGCTTGGACGGGACTGAAGGTATTTTAGAGTTTCTACCC 6960
|||||
Db 6901 ATAAGTGTAGGTCTGGGGCCAGCTTGGACGGGACTGAAGGTATTTTAGAGTTTCTACCC 6960
|||||
Qy 6961 TGTCCCTTCACTTTTCTGAGGAGGAGTCCCTCCCTCAGCTGAGGGGTGGAGTAGGATGG 7020
|||||
Db 6961 TGTCCCTTCACTTTTCTGAGGAGGAGTCCCTCCCTCAGCTGAGGGGTGGAGTAGGATGG 7020
|||||
Qy 7021 TACGTCTGGCTATGGATTTGGCTTTATAAAGAAAGAGGTCTTAAGAAATTTCCCAACC 7080
|||||
Db 7021 TACGTCTGGCTATGGATTTGGCTTTATAAAGAAAGAGGTCTTAAGAAATTTCCCAACC 7080
|||||
Qy 7081 TATGCTTACCTTTTCTGAGCCAGGGGTCTTTGGCTAGGTGGGGGGCTTGGCCCTGTGCC 7140
|||||
Db 7081 TATGCTTACCTTTTCTGAGCCAGGGGTCTTTGGCTAGGTGGGGGGCTTGGCCCTGTGCC 7140
|||||
Qy 7141 TCTCTAAGGGGTGAGTAAGAGACTGTGTGCCCTCCCTTCCCTCCCTCCCTCCAGGGCAG 7200
|||||
Db 7141 TCTCTAAGGGGTGAGTAAGAGACTGTGTGCCCTCCCTTCCCTCCCTCCCTCCAGGGCAG 7200
|||||
Qy 7201 CCAGAGCAAGGTTCCGCCCTCCAGGACCTCTGAGTGTCTTCCGCCACCTGTGCGGATTTCTA 7260
|||||
Db 7201 CCAGAGCAAGGTTCCGCCCTCCAGGACCTCTGAGTGTCTTCCGCCACCTGTGCGGATTTCTA 7260
|||||
Qy 7261 CGAGGGCATCTGCAAAATGGGAGGAGGAGTCCACGCGCTGTGTGACAGTGGGGTGGGC 7320
|||||
Db 7261 CGAGGGCATCTGCAAAATGGGAGGAGGAGTCCACGCGCTGTGTGACAGTGGGGTGGGC 7320
|||||
Qy 7321 CACCTTTCTGTGAGTCCCTTAGGCCGTTTGGAGGACAGGTGAGGACAGCTGCACAGA 7380
|||||
Db 7321 CACCTTTCTGTGAGTCCCTTAGGCCGTTTGGAGGACAGGTGAGGACAGCTGCACAGA 7380
|||||

7381 GGTCTGGGCACTACAGTGGTGTGACAGACCCACGGGCTTGTGAGACTTTCTGGCCAGG 7440
|||||
7381 GGTCTGGGCACTACAGTGGTGTGACAGACCCACGGCTTGTGAGACTTTCTGGCCAGG 7440
|||||
7441 GGCAGCATCTGCCCATCCCTTCGCTGGCGATGGGACTGAGACCCCTCGGGTGGGATGG 7500
|||||
7441 GGCAGCATCTGCCCATCCCTTCGCTGGCGATGGGACTGAGACCCCTCGGGTGGGATGG 7500
|||||
7501 ATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGSCCGCAACCTTGCCTCACTTGC 7560
|||||
7501 ATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGSCCGCAACCTTGCCTCACTTGC 7560
|||||
7561 TCTCCCCACTGSCCCAGGTGCGGCAAGAGTGCATAGTGAAGCCGAGAGCCGAGGGCG 7620
|||||
7561 TCTCCCCACTGSCCCAGGTGCGGCAAGAGTGCATAGTGAAGCCGAGAGCCGAGGGCG 7620
|||||
7621 CCGAGCCGAGAGCCGTGGGGCGAGGAAGCCCGGGAAGCCGCGGGGCCACAGGC 7680
|||||
7621 CCGAGCCGAGAGCCGTGGGGCGAGGAAGCCCGGGAAGCCGCGGGGCCACAGGC 7680
|||||
7681 GGGAGTCCAAGCCAGAGAGCCCGCGCCGCCCAAGAACCAAGCTGGACAGAGGCCTGG 7740
|||||
7681 GGGAGTCCAAGCCAGAGAGCCCGCGCCGCCCAAGAACCAAGCTGGACAGAGGCCTGG 7740
|||||
7741 GCACCGCCAGAGGTGCTGAGGACCCCGCGGAAGCCCTCTGGGACTGTGCTGGCA 7800
|||||
7741 GCACCGCCAGAGGTGCTGAGGACCCCGCGGAAGCCCTCTGGGACTGTGCTGGCA 7800
|||||
7801 CAGCCGAGGCTTGAAGTGGCAGCAGCGCTCAGGTCAGGTCGACACCCGAGCATCACAC 7860
|||||
7801 CAGCCGAGGCTTGAAGTGGCAGCAGCGCTCAGGTCGACACCCGAGCATCACAC 7860
|||||
7861 CCGCGAGGCTTCAAGTGTCTGCTTCCAGAGTGAAGATGAAGGCAAGAGGAGCTGC 7920
|||||
7861 CCGCGAGGCTTCAAGTGTCTGCTTCCAGAGTGAAGATGAAGGCAAGAGGAGCTGC 7920
|||||
7921 TGGTGCCACCAAGATCACTGAGCGCCATCAAGCTGCAACTCAGGSCACAGTGCAGAG 7980
|||||
7921 TGGTGCCACCAAGATCACTGAGCGCCATCAAGCTGCAACTCAGGSCACAGTGCAGAG 7980
|||||
7981 TGCAGATGAAGAGCAGAAAGTGTCAACCCCTAGTACTACACTGTCTTCTTCAAGC 8040
|||||
7981 TGCAGATGAAGAGCAGAAAGTGTCAACCCCTAGTACTACACTGTCTTCTTCAAGC 8040
|||||
8041 GGCAGCGAAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGCTC 8100
|||||
8041 GGCAGCGAAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGCTC 8100
|||||
8101 CGCCTTAGTCCCGCACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGCCCTC 8160
|||||
8101 CGCCTTAGTCCCGCACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGCCCTC 8160
|||||
8161 ACCTTACCCAAACCTAGGTTCCGGTCCCGAGTACAGTGTATCAAAACCCAGATTT 8220
|||||
8161 ACCTTACCCAAACCTAGGTTCCGGTCCCGAGTACAGTGTATCAAAACCCAGATTT 8220
|||||
8221 TCTCAGCTCAAGACCCAGGCTCTGCCCGCTGTTAGAAATATAGTCTCTTCTCCAG 8280
|||||
8221 TCTCAGCTCAAGACCCAGGCTCTGCCCGCTGTTAGAAATATAGTCTCTTCTCCAG 8280
|||||
8281 AATCCAGCCGCGCAATGGAACCTCACGCTGGTCTTAATACCAGTCTTTAAAGGCC 8340
|||||
8281 AATCCAGCCGCGCAATGGAACCTCACGCTGGTCTTAATACCAGTCTTTAAAGGCC 8340
|||||
8341 AGCCCTTAGAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGCAACGTTAC 8400
|||||
8341 AGCCCTTAGAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGCAACGTTAC 8400
|||||
8401 TCAGGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTCAACCTTAAACCAAGGAGCACAG 8460
|||||
8401 TCAGGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTCAACCTTAAACCAAGGAGCACAG 8460
|||||

8461 TCCCACTCCAGCCGGGAGCCTAGGACCACTAGCCCTAGGAGTATATTTCCGCACT 8520
|||||
8461 TCCCACTCCAGCCGGGAGCCTAGGACCACTAGCCCTAGGAGTATATTTCCGCACT 8520
|||||
8521 TCAGAATTCATATCTTGGGAATCCAAAGCTCCTGCCCCAAATTAACCTCAGTCTGCTTC 8580
|||||
8521 TCAGAATTCATATCTTGGGAATCCAAAGCTCCTGCCCCAAATTAACCTCAGTCTGCTTC 8580
|||||
8581 CAGAAATTTGGAATCCTAGTTTCTCTCTGTTATCCGAGTCTGGGACACAAACTCC 8640
|||||
8581 CAGAAATTTGGAATCCTAGTTTCTCTCTGTTATCCGAGTCTGGGACACAAACTCC 8640
|||||
8641 GCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACCTGGCCCGGATC 8700
|||||
8641 GCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACCTGGCCCGGATC 8700
|||||
8701 AGAGCAGACCTCCTCTCGGACCCCTTGGGAACCTCCCAAGAGTCCAGCCCATCTCGGAG 8760
|||||
8701 AGAGCAGACCTCCTCTCGGACCCCTTGGGAACCTCCCAAGAGTCCAGCCCATCTCGGAG 8760
|||||
8761 CATCCCGAGGAAATCTGCAGAGGGTTAGGAGTGGTGCAGAGCCCTGATCTCTCTCT 8820
|||||
8761 CATCCCGAGGAAATCTGCAGAGGGTTAGGAGTGGTGCAGAGCCCTGATCTCTCTCT 8820
|||||
8821 GTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAGATGAATACATTTGTTAAAAAA 8880
|||||
8821 GTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAGATGAATACATTTGTTAAAAAA 8880
|||||
8881 ATATAAGCGCAAGTCCATGTTTATCTGGGAATTTGGGATGGGCGGAGTGGAGCGC 8940
|||||
8881 ATATAAGCGCAAGTCCATGTTTATCTGGGAATTTGGGATGGGCGGAGTGGAGCGC 8940
|||||
8941 CCTCTCTCCCTTTGCTTCTGCTCCCGGACTTTGCGCTCCCTACCTGTGGAGCGCA 9000
|||||
8941 CCTCTCTCCCTTTGCTTCTGCTCCCGGACTTTGCGCTCCCTACCTGTGGAGCGCA 9000
|||||
9001 GGCAGAGTGGCGGCGGAGAGTGGTCCCGCGGCTTGGGCTTCCCGCGCGCG 9060
|||||
9001 GGCAGAGTGGCGGCGGAGAGTGGTCCCGCGGCTTGGGCTTCCCGCGCGCG 9060
|||||
9061 CCGAGGCGGCTTCCCGCGGCGGCTTCCCGGACTGGGCTGGGCGATCCCGGCGCG 9120
|||||
9061 CCGAGGCGGCTTCCCGCGGCGGCTTCCCGGACTGGGCTGGGCGATCCCGGCGCG 9120
|||||
9121 GCCCGCGCGCGGCTTCCAGCCCGCGGCTTCCAGAGCCAGCGGCGCGCGCGCG 9180
|||||
9121 GCCCGCGCGCGGCTTCCAGCCCGCGGCTTCCAGAGCCAGCGGCGCGCGCGCG 9180
|||||

RESULT 2

AC000134/c
LOCUS AC000134 203300 bp DNA PRI 06-MAY-1999
DEFINITION Homo sapiens Chromosome 11q13 BAC Clone 137c7, complete sequence.
AC000134
ACCESSION AC000134 GI:4755212
VERSION
KEYWORDS HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 203300)

AUTHORS

Crabtree, J. and Roe, B.A.

TITLE

Sequence of a 11q13 bac mapping to PYGM

JOURNAL

Unpublished (1997)

REFERENCE

2 (bases 1 to 203300)

AUTHORS

Chandrasekharappa, S.C., Guru, S.C., Manickam, P., Olufemi, S.-E. and Collins, F.S.

TITLE

Sequence of a 11q13 bac mapping to PYGM

JOURNAL

Unpublished (1997)

REFERENCE

3 (bases 1 to 203300)

AUTHORS

Emmert-Buck, M.R., Debelenko, L.V., Zhuang, Z., Lubensky, I.A. and Liotta, L.A.

TITLE

Sequence of a 11q13 bac mapping to PYGM

QY 3300 GTAGAGATGGTGTCTCACTATGTGGCCAGACGTGCTTGAACCTCTTGGGCTCAAGTGAT 3359
Db 160787 GTAGAGATGGTGTCTCACTATGTGGCCAGACGTGCTTGAACCTCTTGGGCTCAAGTGAT 160728
QY 3360 CTGCCGCCCTCAGTCTCCCAAAATGCTGGGATACAGGTGTAGCCACCGCAACTGGTGGC 3419
Db 160727 CTGCCGCCCTCAGTCTCCCAAAATGCTGGGATACAGGTGTAGCCACCGCAACTGGTGGC 160668
QY 3420 CTATGAAAATTTTTTTTTTTTCAGACGCGTCTCACTGTGCGCCAGCGCTGGAGTGC 3479
Db 160667 CTATGAAAATTTTTTTTTTTTCAGACGCGTCTCACTGTGCGCCAGCGCTGGAGTGC 160608
QY 3480 AGTGGTCAATCTCGGCTCACTGCAAGCTCTGCGTCTGCTTTTCATGCCATTTCTCCTGCC 3539
Db 160607 AGTGGTCAATCTCGGCTCACTGCAAGCTCTGCGTCTGCTTTTCATGCCATTTCTCCTGCC 160548
QY 3540 TCCTGCCCTAGCCTCTGAGTGTGGGACTACAGAGCGCTGCCACCATGCCCTGGCTAAT 3599
Db 160547 TCCTGCCCTAGCCTCTGAGTGTGGGACTACAGAGCGCTGCCACCATGCCCTGGCTAAT 160488
QY 3600 TTTTTTTTTGATTTTTTAGTAGACGAGGTTTTACCAATGTTAGCCAGATGGTCTCGATC 3659
Db 160487 TTTTTTTTTGATTTTTTAGTAGACGAGGTTTTACCAATGTTAGCCAGATGGTCTCGATC 160428
QY 3660 TCCTGACCTCGTGATCGCGCGCGCTTGGCCCTCCCAAAGTGTGGGATACAGCGGTGAGC 3719
Db 160427 TCCTGACCTCGTGATCGCGCGCGCTTGGCCCTCCCAAAGTGTGGGATACAGCGGTGAGC 160368
QY 3720 CACGGACCTGGTCAAAAATGTTTGACAGAGAGAGGGGCTTGACCTCAAAAGGCTTAAG 3779
Db 160367 CACGGACCTGGTCAAAAATGTTTGACAGAGAGAGGGGCTTGACCTCAAAAGGCTTAAG 160308
QY 3780 AGTCAGGGCTTGCAAGAGCTTTTGCAACAGCGCGTTGACTGGCAATCCATCCTGGTG 3839
Db 160307 AGTCAGGGCTTGCAAGAGCTTTTGCAACAGCGCGTTGACTGGCAATCCATCCTGGTG 160248
QY 3840 TGCCATATTGAGAAGGAATCAGAGGCTGCTTCTCAGCTTACAGAGGAAAGAGTGCAGAGA 3899
Db 160247 TGCCATATTGAGAAGGAATCAGAGGCTGCTTCTCAGCTTACAGAGGAAAGAGTGCAGAGA 160188
QY 3900 TAAATAGGGTATTCTGCTGGTGTATAGCCAGAGGTGTGGCCAGCGTCTCTGTTT 3959
Db 160187 TAAATAGGGTATTCTGCTGGTGTATAGCCAGAGGTGTGGCCAGCGTCTCTGTTT 160128
QY 3960 TTGCCATTCTGTTTTAACCTAGTAAAGTGCAGTAAATGGAATCCCTTAAATCCATAGAAT 4019
Db 160127 TTGCCATTCTGTTTTAACCTAGTAAAGTGCAGTAAATGGAATCCCTTAAATCCATAGAAT 160068
QY 4020 ATATAATAGAGTTGCAGAGAAAGACGAGGTAGGCCAAAGGCTGGGTACAGGATGATA 4079
Db 160067 ATATAATAGAGTTGCAGAGAAAGACGAGGTAGGCCAAAGGCTGGGTACAGGATGATA 160008
QY 4080 TCCAGAAAGGTATCTCTGTGACATACAGGGTGTAAACAGGGAGAGTCTTTGAACACG 4139
Db 160007 TCCAGAAAGGTATCTCTGTGACATACAGGGTGTAAACAGGGAGAGTCTTTGAACACG 159948
QY 4140 TGGGAGGAAGGATGAGGAGGATAGTGGGAGGAGATCTCAGGTGTGGGTACAGGCTTG 4199
Db 159947 TGGGAGGAAGGATGAGGAGGATAGTGGGAGGAGATCTCAGGTGTGGGTACAGGCTTG 159888
QY 4200 GAAAGGAGTGGGAGGAGTGTGGCCCATCATTACCTGGG-CCCCCTTCCCGCATGTTAAAG 4258
Db 159887 GAAAGGAGTGGGAGGAGTGTGGCCCATCATTACCTGGGCCCCCTTCCCGCATGTTAAAG 159828
QY 4259 CACAGAGACCTCTTTTCAATACCTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 4318
Db 159827 CACAGAGACCTCTTTTCAATACCTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 159768
QY 4319 TGTGGCCCTTTGCTGTGGTGTGGGCTTCCAGGCCCTGGGTCTCGGGATGTCCACCTCGC 4378
Db 159767 TGTGGCCCTTTGCTGTGGTGTGGGCTTCCAGGCCCTGGGTCTCGGGATGTCCACCTCGC 159708
QY 4379 CCTGTCTGAGGATCATGCCCTGGGTAGTGTGTTGGGCCCCAATGGGGAGCAGACAGCTGAGGT 4438

Db 159707 CCTGTCTGAGGATCATGCCCTGGGTAGTGTGTTGGGCCCAATGGGAGCAGACAGCTGAGGT 159648
QY 4439 CACCTGGCAGCGCAAGGGCAACGAGGACCGCAGGGCCACACAGTCAATGCCGTGTGGC 4498
Db 159647 CACCTGGCAGCGCAAGGGCAACGAGGACCGCAGGGCCACACAGTCAATGCCGTGTGGC 159588
QY 4499 TGAGCGGGTATTGTTCCCTCCCGCCAGCGTTGTCCTTCCATCTACTGTAGTCCCAAGCC 4558
Db 159587 TGAGCGGGTATTGTTCCCTCCCGCCAGCGTTGTCCTTCCATCTACTGTAGTCCCAAGCC 159528
QY 4559 ACCCAAGGAGTCCATTTTCTTGGGCGACACCCCTTTCCTCCATCACACCACATPAGG 4618
Db 159527 ACCCAAGGAGTCCATTTTCTTGGGCGACACCCCTTTCCTCCATCACACCACATPAGG 159468
QY 4619 AAGGGRAGACAGAGAGCCCTTTTCTGCTGTCTCATTTCCCTGAAGCAGCAGAGGCTGG 4678
Db 159467 AAGGGRAGACAGAGAGCCCTTTTCTGCTGTCTCATTTCCCTGAAGCAGCAGAGGCTGG 159408
QY 4679 GCCATCATGAGACATATGATCTCATCCCCCCTTAAGAGCTGGCTGTACCTGAAAGGATC 4738
Db 159407 GCCATCATGAGACATATGATCTCATCCCCCCTTAAGAGCTGGCTGTACCTGAAAGGATC 159348
QY 4739 ATACATGCGCTGTGACGCGCAAGATGGAGTGGCGTTTCATGGTGTGCCATCAACCCCTTC 4798
Db 159347 ATACATGCGCTGTGACGCGCAAGATGGAGTGGCGTTTCATGGTGTGCCATCAACCCCTTC 159288
QY 4799 CATTTGACCTGCACACCGACTCGCTGGAGCTTCTGACGCTGCAGCAGGTGAGGCTGAGCC 4858
Db 159287 CATTTGACCTGCACACCGACTCGCTGGAGCTTCTGACGCTGCAGCAGGTGAGGCTGAGCC 159228
QY 4859 AATGGGCGAGGACTGGGCTAGGCGAGACTTTGACTTGTGTGGGACCCCTGGCGAGGSCAC 4918
Db 159227 AATGGGCGAGGACTGGGCTAGGCGAGACTTTGACTTGTGTGGGACCCCTGGCGAGGSCAC 159168
QY 4919 TTTTCCCTTCTGAGCTTCAGCTTCCCTCCTGGAAAATGGGTTAGTAAATTCCTGGGCTG 4978
Db 159167 TTTTCCCTTCTGAGCTTCAGCTTCCCTCCTGGAAAATGGGTTAGTAAATTCCTGGGCTG 159108
QY 4979 GCCTTTCCAGGGCTCTTGGGAGAGTGAATTTGAGATGTGAAATTTGCTTTGACTTCCATTA 5038
Db 159107 GCCTTTCCAGGGCTCTTGGGAGAGTGAATTTGAGATGTGAAATTTGCTTTGACTTCCATTA 159048
QY 5039 AAGGCTGTGTCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGGTTCAGAC 5098
Db 159047 AAGGCTGTGTCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGGTTCAGAC 158988
QY 5099 CCCCACCTCTGCCGATAGGCTAAGGACCGCTTCTCCTCCTGTTCCGTGCTCATAACT 5158
Db 158987 CCCCACCTCTGCCGATAGGCTAAGGACCGCTTCTCCTCCTGTTCCGTGCTCATAACT 158928
QY 5159 CTCTCCTTCCGCTCCTAAGAGCTGCTCTGGCTGCTCTATGACCTTGGGACATCTGGAAGG 5218
Db 158927 CTCTCCTTCCGCTCCTAAGAGCTGCTCTGGCTGCTCTATGACCTTGGGACATCTGGAAGG 158868
QY 5219 TCAGTAGAGGAAGTGGCCAGGCTGGCCCTGGTGGAGCCGGGGGCTGGGTGGCAGGCTG 5278
Db 158867 TCAGTAGAGGAAGTGGCCAGGCTGGCCCTGGTGGAGCCGGGGGCTGGGTGGCAGGCTG 158808
QY 5279 AATTATCATCTTTCTTCCCTAGGTACCCCATGGCTTAGGGAACCTTGGCAGATCTAGAGGAGC 5338
Db 158807 AATTATCATCTTTCTTCCCTAGGTACCCCATGGCTTAGGGAACCTTGGCAGATCTAGAGGAGC 158748
QY 5339 TGGAGCCACCCCTTGGCGGCGCAGACCACTACCCCTTCTACCAAGGTGGGGCATCTA 5398
Db 158747 TGGAGCCACCCCTTGGCGGCGCAGACCACTACCCCTTCTACCAAGGTGGGGCATCTA 158688
QY 5399 AGGAGGTGAGAAAGGAGAGCCCTTAACAGTGGCTGAGCAGGGGCCCTCATCTGGGAGAG 5458
Db 158687 AGGAGGTGAGAAAGGAGAGCCCTTAACAGTGGCTGAGCAGGGGCCCTCATCTGGGAGAG 158628
QY 5459 TGAGAACAGAACTTTTGTGTGTTGGGGGTATCGGCCCATCCAGTCTCACTTTGTGTCAACT 5518

Db 158627 TGAGAGAGAACTTTGTGTGTTGGGGGTATCGCCCATCCAGTCTCACATTTTGTGTCAACT 158568
QY 5519 GTGTGCAGAAATCAGTTCACTCAGGGCTGTCTCAGGGGTGTCCAGGGTTCCTCCAGCCTGGG 5578
Db 158567 GTGTGCAGAAATCAGTTCACTCAGGGCTGTCTCAGGGGTGTCCAGGGTTCCTCCAGCCTGGG 158508
QY 5579 AGTGCAGGGGTGCATTTGTGTCCTCCCTCAGCCTGTCTTTTGTGCCACTGTCTACTGTCTCT 5638
Db 158507 AGTGCAGGGGTGCATTTGTGTCCTCCCTCAGCCTGTCTTTTCTGCCACTGTCTACTGTCTCT 158448
QY 5639 TCCTGGAGTATTAACAGAGGTCAAAATGTGGTAGGACACTGAAGAGGGGTGTCTACTTGG 5698
Db 158447 TCCTGGAGTATTAACAGAGGTCAAAATGTGGTAGGACACTGAAGAGGGGTGTCTACTTGG 158388
QY 5699 TGGGTGTAGTGGGAGAGGGGCCATTTGGGCTGGGCTTGAAGAGTCTTTTGGTGATGTGTAG 5758
Db 158387 TGGGTGTAGTGGGAGAGGGGCCATTTGGGCTGGGCTTGAAGAGTCTTTGGTGATGTGTAG 158328
QY 5759 AAGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCCAGGCCCCACCCCTGCAGTTC 5818
Db 158327 AAGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCCAGGCCCCACCCCTGCAGTTC 158268
QY 5819 TCGCCCGAGCCTCAGCCAGCAGTCTCTGTAGACCCAGGGAGGAGACAGGTAGAAGGGCTG 5878
Db 158267 TCGCCCGAGCCTCAGCCAGCAGTCTCTGTAGACCCAGGGAGGAGACAGGTAGAAGGGCTG 158208
QY 5879 GCAGCGAGTGGAGGTGGGAGTGGAGATGAGAGGACTCCCTGGGATCTTCTCTGTGGCCCT 5938
Db 158207 GCAGCGAGTGGAGGTGGGAGTGGAGATGAGAGGACTCCCTGGGATCTTCTCTGTGGCCCT 158148
QY 5939 TTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGTGCCTCCCTGGA 5998
Db 158147 TTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGTGCCTCCCTGA 158088
QY 5999 GGATCCTGTGCTCACCCTCAGGAGGATTCCTCAGCCAGGACCTACTATCGGGATG 6058
Db 158087 GGATCCTGTGCTCACCCTCAGGAGGATTCCTCAGCCAGGACCTACTATCGGGATG 158028
QY 6059 AACACATCTACCCCTACATGTACCTGGCTGCTACACATGTGCGCAACCGCAATGTGCGGG 6118
Db 158027 AACACATCTACCCCTACATGTACCTGGCTGCTACACATGTGCGCAACCGCAATGTGCGGG 157968
QY 6119 AAGCCCTCGAGCCTGGGGGACAGGGCCACTGTCTATCCAGGAGTGAGGATCCCTCTACT 6178
Db 157967 AAGCCCTCGAGCCTGGGGGACAGGGCCACTGTCTATCCAGGAGTGAGGATCCCTCTACT 157908
QY 6179 AGGGCCTCGAGCCTGTCTTTCTTCTCCCTCCTCATCAGTTTCCAAACACCTCTGTCAGGAC 6238
Db 157907 AGGGCCTCGAGCCTGTCTTTCTTCTCCCTCCTCATCAGTTTCCAAACACCTCTGTCAGGAC 157848
QY 6239 TGAGGCCTGGCTCCAGCGCCCATCCCTTTTCCATCCAGTCCCTAGGCAAGGACAC 6298
Db 157847 TGAGGCCTGGCTCCAGCGCCCATCCCTTTTCCATCCAGTCCCTAGGCAAGGACAC 157788
QY 6299 ATTACCCAGGAGTGGGACCTTGATTAAGGTGTACATCTTTTCCCTCCCTCCCTCTCC 6358
Db 157787 ATTACCCAGGAGTGGGACCTTGATTAAGGTGTACATCTTTTCCCTCCCTCCCTCTCC 157728
QY 6359 TCCTAAATTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACACCATCATCCA 6418
Db 157727 TCCTAAATTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACACCATCATCCA 157668
QY 6419 GCAGTGGGACTTCCACCTTCGCGCCCATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 6478
Db 157667 GCAGTGGGACTTCCACCTTCGCGCCCATCCCTTCTCTCTCTCTCTCTCTCTCTCTCT 157608
QY 6479 GCGTACCCAGACAGCATCATTTTGCAGTGTAGGACCCCACTACTCTCCCAAGCCCTCTGGG 6538
Db 157607 GCGTACCCAGACAGCATCATTTTGCAGTGTAGGACCCCACTACTCTCCCAAGCCCTCTGGG 157548
QY 6539 GCGTCAATCCCGCAGGTCTCTGGGGTACCCCGGATGTGTAGACCCCTTCAGACCTT 6598
Db 157547 GCGTCAATCCCGCAGGTCTCTGGGGTACCCCGGATGTGTAGACCCCTTCAGACCTT 157488

QY 6599 ACAGAGACCCCACTGTCTCTACAGCTTACAACTACTGCCGGGAAGACGAGAGATCTACAA 6658
Db 157487 ACAGAGACCCCACTGTCTCTACAGCTTACAACTACTGCCGGGAAGACGAGAGATCTACAA 157428
QY 6659 GGAGTCTTTTGAAGTAGCCAATGATGTATCCCCAACCTGCTGAAGAGGCGAGCCAGCTT 6718
Db 157427 GGAGTCTTTTGAAGTAGCCAATGATGTATCCCCAACCTGCTGAAGAGGCGAGCCAGCTT 157368
QY 6719 GCTGAGGCGGGGAGGAGCGCGGGGAGCANAAGCCAGGTGAAGGCTGGAGCTCCAG 6778
Db 157367 GCTGAGGCGGGGAGGAGCGCGGGGAGCANAAGCCAGGTGAAGGCTGGAGCTCCAG 157308
QY 6779 CTTGTCTCAGCCTCCCACTGGACAGGCTCCCTTCCACAGGGCCATGGGGCTGCATG 6838
Db 157307 CTTGTCTCAGCCTCCCACTGGACAGGCTCCCTTCCACAGGCTGCATG 157249
QY 6839 ITACGGGATTAGGATGCGCAGGAGGAGGTGGCCCTGAGCAGACAGATATGTTCCTTTTG 6898
Db 157248 ITACGGGATTAGGATGCGCAGGAGGAGGTGGCCCTGAGCAGACAGATATGTTCCTTTTG 157189
QY 6899 CTATAACTCAGGTCTCTGGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGTCTTCTAC 6958
Db 157188 CTATAACTCAGGTCTCTGGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGTCTTCTAC 157129
QY 6959 CTTGTGCCCTTTCAGTTTCATGGCCAGACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGGAT 7018
Db 157128 CTTGTGCCCTTTCAGTTTCATGGCCAGACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGGAT 157069
QY 7019 GGTAGCTCTCGCTATGATGATGCTTTTATAAGGAAGAGGTTCCTAAGAAATGTTCCCAA 7078
Db 157068 GGTAGCTCTCGCTATGATGATGCTTTTATAAGGAAGAGGTTCCTAAGAAATGTTCCCAA 157009
QY 7079 CTTATGCTTACCTTTCTGAGCCAGGGGTCTTTGCTAGGTGGGGGCTTGCCCTGTGC 7138
Db 157008 CTTATGCTTACCTTTCTGAGCCAGGGGTCTTTGCTAGGTGGGGGCTTGCCCTGTGC 156949
QY 7139 CTTCTGCTAAGGGGTGAGTAAGAGACTGATCTGTGCTTCCCTTCCCTTCCCTTCCAGGGC 7198
Db 156948 CTTCTGCTAAGGGGTGAGTAAGAGACTGATCTGTGCTTCCCTTCCCTTCCCTTCCAGGGC 156889
QY 7199 ACCAGAGCAAGGTTCCGCTCCAGGACCTGAGTGTTCGCCCACTGCTGCGCAATTC 7258
Db 156888 ACCAGAGCAAGGTTCCGCTTCCAGGACCTGAGTGTTCGCCCACTGCTGCGCAATTC 156829
QY 7259 TAGCAGGCTCTGCAAAATGGGAGGAGGAGCTCCACCCCTGTGCTGACACCTGGGCTGG 7318
Db 156828 TAGCAGGCTCTGCAAAATGGGAGGAGGAGCTCCACCCCTGTGCTGACACCTGGGCTGG 156769
QY 7319 GGCACCTTTTGTGCTGAGTCCCTTAGGCCCTTTTGAAGGACAGGTGAGGACAGCTGCACA 7378
Db 156768 GGCACCTTTTGTGCTGAGTCCCTTAGGCCCTTTTGAAGGACAGGTGAGGACAGCTGCACA 156709
QY 7379 GAGGTCTGGCACTACAGTGTGTGACAGGACCGGGCTTGTACAGCTTTTCTGCGCCA 7438
Db 156708 GAGGTCTGGCACTACAGTGTGTGACAGGACCGGGCTTGTACAGCTTTTCTGCGCCA 156649
QY 7439 GGGGACAGCATCTGCCATCCCTTCCGTTGCCGATGGACTGAGACCCCTTGGTGGGATG 7498
Db 156648 GGGGACAGCATCTGCCATCCCTTCCGTTGCCGATGGACTGAGACCCCTTGGTGGGATG 156589
QY 7499 GGATGCCAGAGCAGGCTCTGGAGTTCCAGCCACTTGGCCGGCAACCTTGTCTCACTT 7558
Db 156588 GGATGCCAGAGCAGGCTCTGGAGTTCCAGCCACTTGGCCGGCAACCTTGTCTCACTT 156529
QY 7559 GCTCTCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTAGCCGAGAGCCGAGCC 7618
Db 156528 GCTCTCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTAGCCGAGAGCCGAGCC 156469
QY 7619 GSCCGAGGCCGAGGAGCCTGGGGCCAGGAAGCCCGGGAGCCGGCGGGGCCACG 7678
Db 156468 GSCCGAGGCCGAGGAGCCTGGGGCCAGGAAGCCCGGGAGCCGGCGGGGCCACG 156409

Db 155331	GAGCATCCCGAGGAATCTGCAGA -GGGTTAGAGTGGGTGACAAGACCTGATCTCTTT	155273
Qy 8818	CTGTTTTGTACATAGATTTATTTTTCAGTTCACAAGAAAGATGAATACATTTTGTAAAA	8877
Db 155272	CTGTTTTGTACATAGATTTATTTTTCAGTTCACAAGAAAGATCAATTTTGTAAAA	155213
Qy 8878	AAATATATAAGCGCAAGTCATCTTATCTGGAAATTTGGGATGGCGCGGAGTGGAG	8937
Db 155212	AAATATATAAGCGCAAGTCATCTTATCTGGAAATTTGGGATGGCGCGGAGTGGAG	155153
Qy 8938	CGCCCTCTTTCCCTTTTGTCTTCTGGCTCCCGGAGACTTTTCGCTCCCTACCTGTGGAGCG	8997
Db 155152	CGCCCTCTTTCCCTTTTGTCTTCTGGCTCCCGGAGACTTTTCGCTCCCTACCTGTGGAGCG	155093
Qy 8998	CGAGCGACAGTGGCGGCGGAAGACGTAGCTCCGCCCGGCTTGGGGCTTCCCGCCGCG	9057
Db 155092	CGAGCGACAGTGGCGGCGGAAGACGTAGCTCCGCCCGGCTTGGGGCTTCCCGCCGCG	155033
Qy 9058	CCGCCGAGGGCCCTCCCGCGGGCGCTCTCCCGACTGGCGGTGGGGCATCCMGCGG	9117
Db 155032	CCGCCGAGGGCCCTCCCGCGGGCGCTCTCCCGACTGGCGGTGGGGCATCCCGCGG	154974
Qy 9118	CGGCCCGCCCGCCCGGGCTTCAGCCCGCGCCCGCGGCTTCAGAGCCACGGGCGCCCGCC	9177
Db 154973	CGGCCCGCCCGCCCGGGCTTCAGCCCGCGCCCGCGGCTTCAGAGCCACGGGCGCCCGCC	154914
Qy 9178	CCG 9180	
Db 154913	CCG 154911	
RESULT	3	
HS838L14		
LOCUS	HS838L14 109290 bp DNA HTG 31-MAR-1998	
DEFINITION	Homo sapiens chromosome 11 clone 838L14 map 11q13, *** SEQUENCING IN PROGRESS ***, 15 unordered pieces.	
ACCESSION	Y12335	
VERSION	Y12335.1 GI:2208954	
KEYWORDS	HTG; HTGS; PHASE1.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	Kedra,D., Seroussi,E., Fransson,I., Trifunovic,J., Clark,M., Lagercrantz,J., Blennow,E., Mehlin,H. and Dumanski,J.	
TITLE	The germinal center kinase gene and a novel CDC35-like gene are located in the vicinity of the PYGM gene on 11q13	
JOURNAL	Hum. Genet. 100 (5-6), 611-619 (1997)	
MEDLINE	98001089	
REFERENCE	2 (bases 1 to 109290)	
AUTHORS	Kedra,D.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-MAR-1997) D. Kedra, Karolinska Hospital, Department Of Molecular Medicine, Building L-6, S-171 76 Stockholm, SWEDEN	
COMMENT	Warning! This sequence is unfinished. It is not contiguous, consists of 161 contigs separated by gaps of unknown length. Gaps in the sequence are represented by a stretch of NNNN. . . . Order of the contigs unknown.	
	1 2421 contig of 2421	
	2612 4907 contig of 2296	
	5098 5742 contig of 645	
	5933 6485 contig of 553	
	6676 27557 contig of 20882	
	27748 31946 contig of 4199	
	32137 40306 contig of 8170	
	40497 40866 contig of 370	
	41057 49958 contig of 8902	
	50149 53510 contig of 3362	
	57161 57416 contig of 3716	
	57607 58216 contig of 610	

Db 78426 GAGCCCGTGAAGTGCAGGTGAGCCAGATCGAGCCACTGCATTCAGCCCTGGCC 78485
Qy 1361 GAAAGAAAGACCGTGTCTCAAAACAAACAAAGCTACTCTTAGCACGTGTTAGA 1320
Db 78486 GAAAGAAAGACCGTGTCTCAAAACAAACAAAGCTACTCTTAGCACGTGTTAGA 78545
Qy 1321 GTATCTCGCGGGCGAAGTGGGAACAGAGTGTCTGACACAGAGTAGCATCTTTATATGT 1380
Db 78546 GTATCTCGCGGGCGAAGTGGGAACAGAGTGTCTGACACAGAGTAGCATCTTTATATGT 78605
Qy 1381 TAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTCTCTGGCAGTGTAA 1440
Db 78606 TAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTCTCTGGCAGTGTAA 78665
Qy 1441 AAGTATCTGTTTTCTCATATGTGTTTTATTTATTTTCTGGATCAAGCAACCTGATC 1500
Db 78666 AAGTATCTGTTTTCTCATATGTGTTTTATTTATTTTCTGGATCAAGCAACCTGATC 78725
Qy 1501 TTTTTCCTCAACTTCCGACCCGACCGGTGACAGCAAAACCGCAGAAAGCTCGCGGACC 1560
Db 78726 TTTTTCCTCAACTTCCGACCCGACCGGTGACAGCAAAACCGCAGAAAGCTCGCGGACC 78785
Qy 1561 TCCACACCCGAGTCTGACAGTGTGCCCCCGGACTACATTTTCAGAAAGCACTGCGGGC 1620
Db 78786 TCCACACCCGAGTCTGACAGTGTGCCCCCGGACTACATTTTCAGAAAGCACTGCGGGC 78845
Qy 1621 ACGCTTCTGCTGCTGGTGGGCTGAAGGAAGGGCCAAATCCCTGAGTATCTCGGAAGGAG 1680
Db 78846 ACGCTTCTGCTGCTGGTGGGCTGAAGGAAGGGCCAAATCCCTGAGTATCTCGGAAGGAG 78905
Qy 1681 GTGTCGGAGCCGCGGACTAGAGATCCAGAAAGCCAGCGGCGCGGCGCCGCCA 1740
Db 78906 GTGTCGGAGCCGCGGACTAGAGATCCAGAAAGCCAGCGGCGCGGCGCCGCCA 78965
Qy 1741 CTATTTCCAGCTCTCGGGGCGAGGGTGGCCCGAGACTCCACTTCCGGGGGTAGTGC 1800
Db 78966 CTATTTCCAGCTCTCGGGGCGAGGGTGGCCCGAGACTCCACTTCCGGGGGTAGTGC 79025
Qy 1801 GACCTTAGGGGCGGACTTATGTCCAGCAGGCTCGGGGCGGCTGCGCGCGGTGCT 1860
Db 79026 GACCTTAGGGGCGGACTTATGTCCAGCAGGCTCGGGGCGGCTGCGCGCGGTGCT 79085
Qy 1861 AGTGTGGGATGAAGCGGGAGGTGGGAGGGGACCGAGGCGCAGACTCTCTTGGGG 1920
Db 79086 AGTGTGGGATGAAGCGGGAGGTGGGAGGGGACCGAGGCGCAGACTCTCTTGGGG 79145
Qy 1921 TTTGGGGCTTGACCTGGGTGCGCTTTCTGGACAGACTTTTACAGCCCGCGGGGCGACAGT 1980
Db 79146 TTTGGGGCTTGACCTGGGTGCGCTTTCTGGACAGACTTTTACAGCCCGCGGGGCGACAGT 79204
Qy 1981 CGTAGAGAGGGGGGGGCGGCCATTTGGGGCTCCTCAATTGGGGTGTCTTGGGGCGCACCCC 2040
Db 79205 CGTAGAGAGGGGGGGGCGGCCATTTGGGGCTCCTCAATTGGGGTGTCTTGGGGCGCACCCC 79264
Qy 2041 ATCGGGTACCGGGGCTCCCGGAATTGTGGGGGACAAAGGCTCTGAGTCTCGGGTGA 2100
Db 79265 ATCGGGTACCGGGGCTCCCGGAATTGTGGGGGACAAAGGCTCTGAGTCTCGGGTGA 79324
Qy 2101 GGGTCTACCCACAAAGAGGGGAAGCCGGTGAAGAGGCTGAAGAGGGTGGGAAGCA 2160
Db 79325 GGGTCTACCCACAAAGAGGGGAAGCCGGTGAAGAGGCTGAAGAGGGTGGGAAGCA 79384
Qy 2161 GGGAGCTGTGCTGTGTCGGGGGGGTGGAAACCTTACGGGACCTTGGGAGGAGGCTCCC 2220
Db 79385 GGGAGCTGTGCTGTGTCGGGGGGGTGGAAACCTTACGGGACCTTGGGAGGAGGCTCCC 79444
Qy 2221 CGGGCGAACCTTCCCGACCCCTCCCTCCCGGGCTTGCCTTGCAGGCGCGCCACCGCC 2280
Db 79445 CGGGCGAACCTTCCCGACCCCTCCCTCCCGGGCTTGCCTTGCAGGCGCGCCACCGCC 79504
Qy 2281 CGGGCGAACCTTGGGGTGAAGGCCCGCCAGAACGCTGTTCGGCTGCGCTCCATCGAGCA 2340
Db 79504 CGGGCGAACCTTGGGGTGAAGGCCCGCCAGAACGCTGTTCGGCTGCGCTCCATCGAGCA 2340

Db 79505 CGCGCTCATGGGGCTGAAGGCCGCCAGAGACGCTGTTCGCCGTGCGCTCCATCGACGA 79564
Qy 2341 CGTGTGCGCTGTCTGCTGCCAGCTGGGCGGAGAGCGGACCTTGTGCTCCTTTC 2400
Db 79565 CGTGTGCGCTGTCTGCTGCCAGCTGGGCGGAGAGCGGACCTTGTGCTCCTTTC 79624
Qy 2401 CTTGGTGTGGGGCTTCTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGT 2460
Db 79625 CTTGGTGTGGGGCTTCTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGT 79684
Qy 2461 TCCGAGCTCACCTTCCAGCCAGCCCGCCCGAGCCGCTGGCGGCTCACCTACTT 2520
Db 79685 TCCGAGCTCACCTTCCAGCCAGCCCGCCCGAGCCGCTGGCGGCTCACCTACTT 79744
Qy 2521 TCCGCTGGCGGACCTGTCTATCATCGCGGCTCTATGCGCGGCTTACGCGCCAGATCCG 2580
Db 79745 TCCGCTGGCGGACCTGTCTATCATCGCGGCTCTATGCGCGGCTTACGCGCCAGATCCG 79804
Qy 2581 AGCGCGCTGACCTGTCTCTCTATGCTCGAAGAGGGGTGTCTCCAGCCGTGAGCTGGT 2640
Db 79805 AGCGCGCTGACCTGTCTCTCTATGCTCGAAGAGGGGTGTCTCCAGCCGTGAGCTGGT 79864
Qy 2641 GAAGAAGTCTCCGATGTCATATGAACAGCTCAGCGGCTCTACTTCAAGGATCGGGC 2700
Db 79865 GAAGAAGTCTCCGATGTCATATGAACAGCTCAGCGGCTCTACTTCAAGGATCGGGC 79924
Qy 2701 CCACATCCAGTCCCTCTCTAGCTTATCATACAGGTTGGAGCCAGTAGGTGGGAATCTTAT 2760
Db 79925 CCACATCCAGTCCCTCTCTAGCTTATCATACAGGTTGGAGCCAGTAGGTGGGAATCTTAT 79984
Qy 2761 CCATGACCCACTTCTTCAAAACCTTCCATGTTTACAGAACCTTTTAAAGAACTGTAAAGC 2820
Db 79985 CCATGACCCACTTCTTCAAAACCTTCCATGTTTACAGAACCTTTTAAAGAACTGTAAAGC 80044
Qy 2821 CTTCTGAGGTTGCGCAGGTGTATTTTCTCTCTTGTGAGTTGGGAACTGAAGCCAGAGA 2880
Db 80045 CTTCTGAGGTTGCGCAGGTGTATTTTCTCTCTTGTGAGTTGGGAACTGAAGCCAGAGA 80104
Qy 2881 GGGAAATGATATGCCAAAGTCACACGCGCATGGCAGGCTGGAAGCTGAAGCTTGAATCA 2940
Db 80105 GGGAAATGATATGCCAAAGTCACACGCGCATGGCAGGCTGGAAGCTGAAGCTTGAATCA 80164
Qy 2941 CTTGGCTCCAAATCATCAACCTCAGCTTGCCTCTCAGCACCCGCCCTTGGCAGCTGA 3000
Db 80165 CTTGGCTCCAAATCATCAACCTCAGCTTGCCTCTCAGCA-CCCCAGCTTGGCAGCTGA 80223
Qy 3001 ACAGCTACAGGAGTCTTAAGCATGAGACAGAGGGCGGACAGAGTTAGGGGCAAGA 3060
Db 80224 ACAGCTACAGGAGTCTTAAGCATGAGACAGAGGGCGGACAGAGTTAGGGGCAAGA 80283
Qy 3061 AGATGAAATGGGCTGCAATTTGAGGCAAGTTAAACAAATAATGGCTATGAAGATTTTTTT 3120
Db 80284 AGATGAAATGGGCTGCAATTTGAGGCAAGTTAAACAAATAATGGCTATGAAGATTTTTTT 80343
Qy 3121 TTTTTTTTTTTTTTTGGACAGAGGCTCAGCTGTCTCCCGCAGGCTGGAGTGGAGTGGTG 3180
Db 80344 TTTTTTTTTTTTTTTGGACAGAGGCTCAGCTGTCTCCCGCAGGCTGGAGTGGAGTGGTG 80403
Qy 3181 ATCATGGCTCAGCTCAGCTCCTCGGCTCAGATCCCTCGGCTCAGATCTCCAACTCAGCTCC 3240
Db 80404 ATCATGGCTCAGCTCAGCTCCTCGGCTCAGATCTCCCTCGGCTCAGATCTCCAACTCAGCTCC 80463
Qy 3241 TGAGTAGCTGAGAGTACAGGATGACCGGTGTGCTGTTTAAATTTTTTGTATTTTTTTTGG 3300
Db 80464 TGAGTAGCTGAGAGTACAGGATGACCGGTGTGCTGTTTAAATTTTTTGTATTTTTTTTGG 80523
Qy 3301 TAGAGATGGTGTCTCAGTATGTGGCCAGAGTGGTCTTGAAGTCTTGGGCTCAAGTGTATC 3360
Db 80524 TAGAGATGGTGTCTCAGTATGTGGCCAGAGTGGTCTTGAAGTCTTGGGCTCAAGTGTATC 80583
Qy 3361 TGCCCGCTCAGTCTCCCAATGTGGGATTTACAGGCTGTGAGCCACCGCAACTGGTGGCC 3420
Db 80584 TGCCCGCTCAGTCTCCCAATGTGGGATTTACAGGCTGTGAGCCACCGCAACTGGTGGCC 80643

Db 82804 GTGCGAGGGCTGCTATTGTTCCCTCAGCCCTGCTTTCTGCCACTGCTTACTGTCCTT 82863
Qy 5640 CTTGGAGTATAACAGAGGTCAATGTGGTAGGACACTGAAGAGGGGTGTTCACCTGGT 5699
Db 82864 CTTGGAGTATAACAGAGGTCAATGTGGTAGGACACTGAAGAGGGGTGTTCACCTGGT 82923
Qy 5700 GGGTGTAGGTGGGAGGAGGCCATTTGGGCTGGGCTTTGAAAGTCTTTGGTGATGTGAGA 5759
Db 82924 GGGTGTAGGTGGGAGGAGGCCATTTGGGCTGGGCTTTGAAAGTCTTTGGTGATGTGAGA 82983
Qy 5760 AGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGAGGGCCACCCCTCGAGTCT 5819
Db 82984 AGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGAGGGCCACCCCTCGAGTCT 83043
Qy 5820 GCCCCAGGCCCTCAGCCAGCAGCTCTGTAGACCCAGGAGGAGAGACAGGTAGAAGGGCTGG 5879
Db 83044 GCCCCAGGCCCTCAGCCAGCAGCTCTGTAGACCCAGGAGGAGAGACAGGTAGAAGGGCTGG 83103
Qy 5880 CAGGAGTGGAGGTGGGAGTGGAGTGGAGAGGAGTCCCTGGGATCTTCTGTGGCCCT 5939
Db 83104 CAGGAGTGGAGGTGGGAGTGGAGTGGAGAGGAGTCCCTGGGATCTTCTGTGGCCCT 83163
Qy 5940 TCTGGGTGTGCCCTGTGGGGGCAATTTGTGCCAGAGGCACTGGGGTGGCTCCCTGAG 5999
Db 83164 TCTGGGTGTGCCCTGTGGGGGCAATTTGTGCCAGAGGCACTGGGGTGGCTCCCTGAG 83223
Qy 6000 GATCCTCTGCCCTCACCCTCATTCCAGGCAATTCCTCAGCCAAAGACCTACTATCGGGATGA 6059
Db 83224 GATCCTCTGCCCTCACCCTCATTCCAGGCAATTTGCCCTCAGCCAAAGACCTACTATCGGGATGA 83283
Qy 6060 ACACATCTACCCCTACATGTACCTGGCTGCTACACTGTCCGCAACCGCAATGTGCGGGA 6119
Db 83284 ACACATCTACCCCTACATGTACCTGGCTGCTACACTGTCCGCAACCGCAATGTGCGGGA 83343
Qy 6120 AGCCCTGAGGCTGTGGGGGACAGGCGCACTGTATCCAGAGGTGAGGATCCCCCTACTA 6179
Db 83344 AGCCCTGAGGCTGTGGGGGACAGGCGCACTGTATCCAGAGGTGAGGATCCCCCTACTA 83403
Qy 6180 GGGCTCGAGCCTGCTCTTCTTCTCCCTCCATCATAGTTTCCAAACACCCCTCGTCCAGACT 6239
Db 83404 GGGCTCGAGCCTGCTCTTCTTCTCCCTCCATCATAGTTTCCAAACACCCCTCGTCCAGACT 83463
Qy 6240 GAGGCTGGCTCCGAGCCGCACTCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCA 6299
Db 83464 GAGGCTGGCTCCGAGCCGCACTCCCTTTCCATCCGCTCCCTAGGAGCAAGGCCACCA 83523
Qy 6300 TTACCCAGGAGGTAGGAGCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCCT 6359
Db 83524 TTACCCAGGAGGTAGGAGCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCCT 83583
Qy 6360 CCTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAATGTTTAAACCACCATCATCCAG 6419
Db 83584 CCTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAATGTTTAAACCACCATCATCCAG 83643
Qy 6420 CAGTGGGACTTCCACCTCGGCCCACTGCCCTCCCTCATCTTGTCTTCTCCTCTGG 6479
Db 83644 CAGTGGGACTTCCACCTCGGCCCACTGCCCTCCCTCATCTTGTCTTCTCCTCTGG 83703
Qy 6480 CTTGACCCAGACAGCATTTTGTGAGTGGAGCCCACTTCTCCCTCCCTCCCTCCCTGGG 6539
Db 83704 CTTGACCCAGACAGCATTTTGTGAGTGGAGCCCACTTCTCCCTCCCTCCCTCCCTGGG 83763
Qy 6540 GCTCATCCCCCGGAGGTCTGGGGTACCCCGGTGAGTGTGAGACCCCTTTCAGACCCCTA 6599
Db 83764 GCTCATCCCCCGGAGGTCTGGGGTACCCCGGTGAGTGTGAGACCCCTTTCAGACCCCTA 83823
Qy 6600 CAGAGACCCCACTGCTCTCAGAGTACAACTACTGCCGGGAGAGGAGATCTACAAG 6659
Db 83824 CAGAGACCCCACTGCTCTCAGAGTACAACTACTGCCGGGAGAGGAGATCTACAAG 83883
Qy 6660 GAGTCTTTGAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGAGGAGGCTTG 6719
Db 83984 GAGTCTTTGAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGAGGAGGCTTG 85022

Db 83984 GAGTCTTTGAAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGAGGAGGCTTG 83943
Qy 6720 CTGAGGCGGGCAGGAGCGGCCGGGAGCAAAAGCCAGGTGAAGGCTGAGGCTTCCAGC 6779
Db 83944 CTGAGGCGGGCAGGAGCGGCCGGGAGCAAAAGCCAGGTGAAGGCTGAGGCTTCCAGC 84003
Qy 6780 CTGTGTCCAGCCTCCCACTTGAGCAGGGCTCCCTTCCACAGGGCATGGGGCTGCATCT 8439
Db 84004 CTGTGTCCAGCCTCCCACTTGAGCAGGGCTCCCTTCCACA -GGCCATGGGGCTGCATGT 84062
Qy 6840 ACGGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGC 6899
Db 84063 ACGGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGC 84122
Qy 6900 TATAACTGAGGTCTCGGCCCCACCTTGGACGGGACTGAAGCTATTTAGAGCTTTCTACC 6959
Db 84123 TATAACTGAGGTCTCGGCCCCACCTTGGACGGGACTGAAGCTATTTAGAGCTTTCTACC 84182
Qy 6960 CTGTGCTTACAGTTTCAATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGATG 7019
Db 84183 CTGTGCTTACAGTTTCAATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGATG 84242
Qy 7020 GTACGTCTGCTATGATTTGGCTTTTATAAAGGAAAGAGGTTCTAAGAAATGTTCCCAAC 7079
Db 84243 GTACGTCTGCTATGATTTGGCTTTTATAAAGGAAAGAGGTTCTAAGAAATGTTCCCAAC 84302
Qy 7080 CTATGCTTACCTTTTCTGGAGCCAGGGCTTTGCTAGGTGGGGGCTGCGCTGTGCC 7139
Db 84303 CTATGCTTACCTTTTCTGGAGCCAGGGCTTTGCTAGGTGGGGGCTGCGCTGTGCC 84362
Qy 7140 CTCTGCTTAAGGGGTGAGTAAAGAGACTGATCTGTGCTCCCTTCCCTCCCTCCCTCAGGSCA 7199
Db 84363 CTCTGCTTAAGGGGTGAGTAAAGAGACTGATCTGTGCTCCCTTCCCTCCCTCAGGSCA 84422
Qy 7200 CCCAGAGCCAAAGTTTCCGCCCTCCAGGACCTGAGTGTCTCGGCCACCTGCTGGGATCT 7259
Db 84423 CCCAGAGCCAAAGTTTCCGCCCTCCAGGACCTGAGTGTCTCGGCCACCTGCTGGGATCT 84482
Qy 7260 ACAGCGCATCTCGAAATGGGAGGAGGAGTCCACAGCCCTGCTGACCTGGGCTGGG 7319
Db 84483 ACAGCGCATCTCGAAATGGGAGGAGGAGTCCACAGCCCTGCTGACCTGGGCTGGG 84542
Qy 7320 CCACCTTTCTGTGCTCAGTCCCTAGGCGGTGTTGAGGGACAGGTGAGGACAGCTGCACAG 7379
Db 84543 CCACCTTTCTGTGCTCAGTCCCTAGGCGGTGTTGAGGGACAGGTGAGGACAGCTGCACAG 84602
Qy 7380 AGGTCTGGGCACTACAGGTGTGACAGCAGCCAGGGCTTGTACAGACTTTTCTGCCACAG 7439
Db 84603 AGGTCTGGGCACTACAGGTGTGACAGCAGCCAGGGCTTGTACAGACTTTTCTGGCCACAG 84662
Qy 7440 GGGCAGCATCTGCCCATCCCTTCCGGTCCGATGGGACTGAGACCCCTGGGTGGGATGG 7499
Db 84663 GGGCAGCATCTGCCCATCCCTTCCGGTCCGATGGGACTGAGACCCCTGGGTGGGATGG 84722
Qy 7500 GATGGCCAGACAGGGTCTGGAGTTCAGCCACTTGGCCGCAACCTTGTCTCACCTTG 7559
Db 84723 GATGGCCAGACAGGGTCTGGAGTTCAGCCACTTGGCCGCAACCTTGTCTCACCTTG 84782
Qy 7560 CTCTCCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTGAGCCGAGAGCCGAGCGG 7619
Db 84783 CTCTCCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTGAGCCGAGAGCCGAGCGG 84842
Qy 7620 GCCGAGCCGAGGAGCCGTGGGGGAGGAAGCCCGGAGGAGCCGCGGGGCGCCACCGG 7679
Db 84843 GCCGAGCCGAGGAGCCGTGGGGGAGGAAGCCCGGAGGAGCCGCGGGGCGCCACCGG 84902
Qy 7680 CGGAGTCCAAAGCCAGAGGAGCCCGCCCGCCCAAGAGCCAGCACTGGCAAGGCGCTG 7739
Db 84903 CGGAGTCCAAAGCCAGAGGAGCCCGCCCGCCCAAGAGCCAGCACTGGCAAGGCGCTG 84952
Qy 7740 GGCACCGCCAGGAGTGTGAGTGTAGGACCCCGCCGAGAGCCCTCCTGGGACTGTGCTCTGC 7799
Db 84963 GGCACCGCCAGGAGTGTGAGTGTAGGACCCCGCCGAGAGCCCTCCTGGGACTGTGCTCTGC 85022

```
QY 7800 ACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCGACGCCCGCAGCATCACCA 7859
|||||
Db 85023 ACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCGACGCCCGCAGCATCACCA 85082
|||||
QY 7860 CGCCCGAGGCTCCAGTGTCTACTTCCAGAGTGAGAGATGAAGGGCATGAAGAGAGCTG 7919
|||||
Db 85083 CGCCCGAGGCTCCAGTGTCTACTTCCAGAGTGAGAGATGAAGGGCATGAAGAGAGCTG 85142
|||||
QY 7920 CTGGTGGCCACCAAGATCAACTCGAGCGCATCAAGTCAACTGCTGCTTTCCTCAAG 7979
|||||
Db 85143 CTGGTGGCCACCAAGATCAACTCGAGCGCATCAAGTCAACTGCTGCTTTCCTCAAG 85202
|||||
QY 7980 GTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTGCTGCTTTCCTCAAG 8039
|||||
Db 85203 GTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTGCTGCTTTCCTCAAG 85262
|||||
QY 8040 CGGCAGCGCAAGGCGCTTGAAGTCTGAGTCTGCGGACTTCGGACCGCTTGTGGGACCCAGGCT 8099
|||||
Db 85263 CGGCAGCGCAAGGCGCTTGAAGTCTGAGTCTGCGGACTTCGGACCGCTTGTGGGACCCAGGCT 85322
|||||
QY 8100 CGGCCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCT 8159
|||||
Db 85323 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85382
|||||
QY 8160 CACCTCTACCCAAACCTAGTTCGCGTCCCGAGTACAGTCTGTATCAAAACCCAGCAT 8219
|||||
Db 85383 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85442
|||||
QY 8220 TTCCTCAGCTCAGAACCCAGGCTCTGCCCCAGTCTGTTAGATATAGTCTCTT 8273
|||||
Db 85443 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 85496
|||||

RESULT 4
AP000928/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-772K10 map 11q13, WORKING
DRAFT SEQUENCE, 33 unordered pieces.
ACCESSION
AP000928.2 GI:8119069
VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens DNA, clone:RP11-772K10.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 171980)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 171,980 genomic DNA of 11q13
Published Only in Database (1999) In press
REFERENCE
2 (bases 1 to 171980)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamiara, Kanagawa 228-8555,
Japan [E-mail:hattori@gs.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924]
On May 31, 2000 this sequence version replaced gi:6997744.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gs.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-772K10
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
```

Assembly program: Phrap; version 0.990329

Consensus quality: 141099 bases at least Q40

Consensus quality: 154862 bases at least Q30

Consensus quality: 163556 bases at least Q20

Insert size: 168780; sum-of-contigs

Quality coverage: 3.98x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 13598 contig of 13598 bp in length
26815 contig of 13117 bp in length
37862 contig of 10947 bp in length
46388 contig of 8426 bp in length
55060 contig of 8572 bp in length
61498 contig of 6338 bp in length
68961 contig of 7363 bp in length
76353 contig of 7292 bp in length
84390 contig of 7937 bp in length
91104 contig of 6614 bp in length
97347 contig of 6143 bp in length
103810 contig of 6363 bp in length
110422 contig of 6512 bp in length
116801 contig of 6279 bp in length
121989 contig of 5088 bp in length
125946 contig of 3857 bp in length
132109 contig of 6063 bp in length
136422 contig of 4213 bp in length
140566 contig of 4044 bp in length
144574 contig of 3908 bp in length
148311 contig of 3637 bp in length
150652 contig of 2241 bp in length
153335 contig of 2583 bp in length
155668 contig of 2233 bp in length
158851 contig of 3083 bp in length
161241 contig of 2290 bp in length
163307 contig of 1966 bp in length
163408 contig of 2080 bp in length
167734 contig of 2147 bp in length
168855 contig of 1021 bp in length
167835 contig of 1134 bp in length
168956 contig of 1134 bp in length
170089 contig of 466 bp in length
170190 contig of 466 bp in length
170756 contig of 1225 bp in length
171980 contig of 1225 bp in length
```

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 13598 contig of 13598 bp in length
13599 13598 gap of 100 bp
13699 26815 contig of 13117 bp in length
26816 26915 gap of 100 bp
26916 37862 contig of 10947 bp in length
37863 37962 gap of 100 bp
37963 46388 contig of 8426 bp in length
46389 46488 gap of 100 bp
46489 55060 contig of 8572 bp in length
55061 55160 gap of 100 bp
55161 61498 contig of 6338 bp in length
61499 61598 gap of 100 bp
61599 68961 contig of 7363 bp in length
68962 69061 gap of 100 bp
69062 76353 contig of 7292 bp in length
76354 76453 gap of 100 bp
76454 84390 contig of 7937 bp in length
```

```
* 84391 84490: gap of 100 bp
* 84491 91104: contig of 6614 bp in length
* 91105 91204: gap of 100 bp
* 91205 97347: contig of 6143 bp in length
* 97348 97447: gap of 100 bp
* 97448 103810: contig of 6363 bp in length
* 103811 103910: gap of 100 bp
* 103911 110422: contig of 6512 bp in length
* 110423 110522: gap of 100 bp
* 110523 116801: contig of 6279 bp in length
* 116802 116901: gap of 100 bp
* 116902 121989: contig of 5088 bp in length
* 121990 122089: gap of 100 bp
* 122090 125946: contig of 3857 bp in length
* 125947 126046: gap of 100 bp
* 126047 132109: contig of 6063 bp in length
* 132110 132209: gap of 100 bp
* 132210 136422: contig of 4213 bp in length
* 136423 136522: gap of 100 bp
* 136523 140566: contig of 4044 bp in length
* 140567 140666: gap of 100 bp
* 140667 144574: contig of 3908 bp in length
* 144575 144674: gap of 100 bp
* 144675 148311: contig of 3637 bp in length
* 148312 148411: gap of 100 bp
* 148412 150652: contig of 2241 bp in length
* 150653 150752: gap of 100 bp
* 150753 153335: contig of 2583 bp in length
* 153336 153435: gap of 100 bp
* 153436 155668: contig of 2233 bp in length
* 155669 155768: gap of 100 bp
* 155769 158851: contig of 3083 bp in length
* 158852 158951: gap of 100 bp
* 158952 161241: contig of 2290 bp in length
* 161242 161341: gap of 100 bp
* 161342 163307: contig of 1966 bp in length
* 163308 163407: gap of 100 bp
* 163408 165487: contig of 2080 bp in length
* 165488 165587: gap of 100 bp
* 165588 167734: contig of 2147 bp in length
* 167735 167834: gap of 100 bp
* 167835 168855: contig of 1021 bp in length
* 168856 168955: gap of 100 bp
* 168956 170089: contig of 1134 bp in length
* 170090 170189: gap of 100 bp
* 170190 170655: contig of 466 bp in length
* 170656 170755: gap of 100 bp
* 170756 171980: contig of 1225 bp in length.
```

FEATURES

source

```
1. .171980
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="RP11-772K10"
1. .13598
/feature="assembly_fragment"
13699. .26815
/feature="assembly_fragment"
26916. .37862
/feature="assembly_fragment"
37963. .46388
/feature="assembly_fragment"
46489. .55060
/feature="assembly_fragment"
55161. .61498
/feature="assembly_fragment"
61599. .68961
/feature="assembly_fragment"
69062. .76353
/feature="assembly_fragment"
76454. .84390
/feature="assembly_fragment"
```

```
misc_feature 84491. .91104
/feature="assembly_fragment"
misc_feature 91205. .97347
/feature="assembly_fragment"
misc_feature 97448. .103810
/feature="assembly_fragment"
misc_feature 103911. .110422
/feature="assembly_fragment"
misc_feature 110523. .116801
/feature="assembly_fragment"
misc_feature 116902. .121989
/feature="assembly_fragment"
misc_feature 122090. .125946
/feature="assembly_fragment"
misc_feature 126047. .132109
/feature="assembly_fragment"
misc_feature 132210. .136422
```

Query Match 63.7%; Score 5844; DB 82; Length 171980;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 6036; Conservative 0; Mismatches 120; Indels 20; Gaps 16;

```
QY 2421 AGCATTTTCGGCTGTCAACCGCGTCATCCCTACCAACGTTCCCGAGCTCACCTTCAGC 2480
DB 132109 AGCATTTTCGGCTGTCAACCGCGTCATCCCTACCAACGTTCCCGAGCTCACCTTCAGC 132050
QY 2481 CCAGCCCCCGCCCGAGCCGCTGGCGGCTTCACCTACTTTCCCGTGGCGGACCTGTCTA 2540
DB 132049 CCAGCCCCCGCCCGAGCCGCTGGCGGCTTCACCTACTTTCCCGTGGCGGACCTGTCTA 131990
QY 2541 TCATCGCGCGCTCTATGCGCGCTTCACCGCCAGATCCGAGCGCGCTGCACCTGTCCC 2600
DB 131989 TCATCGCGCGCTCTATGCGCGCTTCACCGCCAGATCCGAGCGCGCTGCACCTGTCCC 131930
QY 2601 TCTATCTCGAAGAGGGGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCCTCCATGTCA 2660
DB 131929 TCTATCTCGAAGAGGGGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCCTCCATGTCA 131870
QY 2661 TATGGAACAGCTCAGCGCTCTACTTCAAGAGTCCGGGCGCACATCCAGTCCCTCTTCA 2720
DB 131869 TATGGAACAGCTCAGCGCTCTACTTCAAGAGTCCGGGCGCACATCCAGTCCCTCTTCA 131810
QY 2721 GCTTCATCAGAGTTGGAGCCAGTAGTGGAAATCTATTCCATGACCCACTTCTTCAAA 2780
DB 131809 GCTTCATCAGAGTTGGAGCCAGTAGTGGAAATCTATTCCATGACCCACTTCTTCAAA 131750
QY 2781 ACCCTCATGTTTACAGACCCCTTTTAAAGAACTGTAAAGCTTGTAGGTTTGGCAGGTG 2840
DB 131749 ACCCTCATGTTTACAGACCCCTTTTAAAGAACTGTAAAGCTTGTAGGTTTGGCAGGTG 131690
QY 2841 TTATTTTCTCTTTTCAGAGTTGGGAACTGAAGCCCGAGAGGGGAAATGATATGCCAAG 2900
DB 131689 TTATTTTCTCTTTTCAGAGTTGGGAACTGAAGCCCGAGAGGGGAAATGATATGCCAAG 131630
QY 2901 TCACACAGGCGATGGCAGGGCTGGAAGTGAAGCCGTGATCATCTTGGCTCCCAATCATCAAC 2960
DB 131629 TCACACAGGCGATGGCAGGGCTGGAAGTGAAGCCGTGATCATCTTGGCTCCCAATCATCAAC 131570
QY 2961 CTCACCTCTGCCCTCTCAGCACCCCTTGGCAGCTGAACAGCTACAGGAGTTCTTAAG 3020
DB 131569 CTCACCTCTGCCCTCTCAGCACCCCTTGGCAGCTGAACAGCTACAGGAGTTCTTAAG 131510
QY 3021 CATGAGACACAGGGGGCGGACGAGATTTAGGGGCGAAGAGATCAAAATGGGCTGCATT 3080
DB 131509 CATGAGACACAGGGGGCGGACGAGATTTAGGGGCGAAGAGATCAAAATGGGCTGCATT 131450
QY 3081 TGAGGCGAGTTAAACAAAATATGCGTATGAAGATTTTTTTTTTTTTTTTTTTTGTGAGACA 3140
DB 131449 TGAGGCGAGTTAAACAAAATATGCGTATGAAGATTTTTTTTTTTTTTTTTTTTGTGAGACA 131391
QY 3141 GGGTCTCAGCTCTGCCCCAGGCTGGAGTGCAGTGGTGTGATCATGGCTCACTCAGCGCT 3200
DB 131390 GGGTCTCAGCTCTGCCCCAGGCTGGAGTGCAGTGGTGTGATCATGGCTCACTCAGCGCT 131331
```


[illegible]

RESULT 5
AP000575/c
LOCUS

Db 17886 GGCAACAGGAGCCGAGGGCCACAGACGATCAATGCCGGTGTGGCTGAGGGGTATTGTT 17827
QY 4514 CCCTCCCCCAGCCTTGCTCCCTTTCTACTACTGTAGTCCCAAGCCACCCAGGACTCCA 4573
Db 17826 CCCTCCCCCAGCCTTGCTCCCTTTCTACTACTGTAGTCCCAAGCCACCCAGGACTCCA 17767
QY 4574 TTTTCTTGGGCCACACCCCTTTCTTCCCATCACACCCACATAGGAAGGAAGACAAAG 4633
Db 17766 TTTTCTTGGGCCACACCCCTTTCTTCCCATCACACCCACATAGGAAGGAAGACAAAG 17707
QY 4634 AGCCCTTTTCTGGCTGTCTATCCCTGAAGCAGGACAGAGGTGGGCCATCATGAGACAT 4693
Db 17706 AGCCCTTTTCTGGCTGTCTATCCCTGAAGCAGGACAGAGGTGGGCCATCATGAGACAT 17647
QY 4694 AATGATCTCATCCCCCTTAAGAGCTGGCTGTACCTGAAGAGGATCATACATGCGCTGA 4753
Db 17646 AATGATCTCATCCCCCTTAAGAGCTGGCTGTACCTGAAGAGGATCATACATGCGCTGA 17587
QY 4754 CCACAAGATGGAGTGGCTTTCATGTGTGTGCCATCAACCCCTTCCATTTGACCTGCAAC 4813
Db 17586 CCACAAGATGGAGTGGCTTTCATGTGTGTGCCATCAACCCCTTCCATTTGACCTGCAAC 17527
QY 4814 CGACTCGCTGGAGTCTGACAGCTGCAGAGGTGAGGGCTGAGGCAATGGGGCAGGACTG 4873
Db 17526 CGACTCGCTGGAGTCTGACAGCTGCAGAGGTGAGGGCTGAGGCAATGGGGCAGGACTG 17467
QY 4874 GGCTAGGCCAGACTTGACTTGTCTGGGACCCCTGGGCAGGGGACATTTCCCTTCCCTGAGC 4933
Db 17466 GGCTAGGCCAGACTTGACTTGTCTGGGACCCCTGGGCAGGGGACATTTCCCTTCCCTGAGC 17407
QY 4934 TTGAGCTTCCCTTCTGGAAAAATGGGTAGTAAATTCCTTGGCCCTGGCTTTTCCAGGGCT 4993
Db 17406 TTGAGCTTCCCTTCTGGAAAAATGGGTAGTAAATTCCTTGGCCCTGGCTTTTCCAGGGCT 17347
QY 4994 CTGGGAGAGTAGAATGAGATGTGAATTCCTTGGACTCCATTAAGAGGCTGGTCCAG 5053
Db 17346 CTGGGAGAGTAGAATGAGATGTGAATTCCTTGGACTCCATTAAGAGGCTGGTCCAG 17287
QY 5054 AATTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGGTTCTGACCCCCACCTCTGGCCG 5113
Db 17286 AATTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGGTTCTGACCCCCACCTCTGGCCG 17227
QY 5114 ATAGGCTAAGACCCGTTCTCCCTGTTCCGTTGGCTCATAACTCTCTCTTGGGTCC 5173
Db 17226 ATAGGCTAAGACCCGTTCTCCCTGTTCCGTTGGCTCATAACTCTCTCTTGGGTCC 17167
QY 5174 TAGAAGCTGCTGGCTGTCTATGACCTGGGACATCTGAAAGGTCAGTAGAGGGAAT 5233
Db 17166 TAGAAGCTGCTGGCTGTCTATGACCTGGGACATCTGAAAGGTCAGTAGAGGGAAT 17107
QY 5234 GGCCAGGCTGGCTGGTGGGCGGGGGCTGGGTGGCAGCCTGAATTTATGATCTTTC 5293
Db 17106 GGCCAGGCTGGCTGGTGGGCGGGGGCTGGGTGGCAGCCTGAATTTATGATCTTTC 17047
QY 5294 CTAGGTACCCCATGGCCTTAGGAAACCTGCGAGATCTAGAGAGCTGGAGCCCAACCCCTG 5353
Db 17046 CTAGGTACCCCATGGCCTTAGGAAACCTGCGAGATCTAGAGAGCTGGAGCCCAACCCCTG 16987
QY 5354 GCCGGCCAGACCCACTCACCTCTACACAAGGTGGGGGATCTAAGAGAGGTGCACAAG 5413
Db 16986 GCCGGCCAGACCCACTCACCTCTACACAAGGTGGGGGATCTAAGAGAGGTGCACAAG 16927
QY 5414 GGAGACCTTAACAGTGGCTGAGGCAGGGGCCCTCATCTGGGCAGATGAGAAGAACTTT 5473
Db 16926 GGAGACCTTAACAGTGGCTGAGGCAGGGGCCCTCATCTGGGCAGATGAGAAGAACTTT 16867
QY 5474 GTGTGTGGGGGTAAAGCCCATCCAGTCTACCTTTGTGCAACTGTGTGCAAGATCAGT 5533
Db 16866 GTGTGTGGGGGTAAAGCCCATCCAGTCTACCTTTGTGCAACTGTGTGCAAGATCAGT 16807
QY 5534 TCAGTACAGGCTGTCTAGAGGGTGTCCAGGGTTCACAGGCTGGGAGTGGCAGGGGTGC 5593
Db 16806 TCAGTACAGGCTGTCTAGAGGGTGTCCAGGGTTCACAGGCTGGGAGTGGCAGGGGTGC 16747

QY 5594 ATTTGTCCCTCAGCCCTGCTTTTCTTCCACATGCTTACTGTCTCTCTCTGGAGTATAACA 5653
Db 16746 ATTTGTCCCTCAGCCCTGCTTTTCTTCCACATGCTTACTGTCTCTCTGGAGTATAACA 16687
QY 5654 GAGTCAAAATGTGTAGGAGCACTGAAGAGGGGTGTTCACCTTGGTGGGTGTAGTGGGG 5713
Db 16686 GAGTCAAAATGTGTAGGAGCACTGAAGAGGGGTGTTCACCTTGGTGGGTGTAGTGGGG 16627
QY 5714 AGAGGGCCATTTGGCTTGGCTTGAAGTCTTTGGTGATGTGTAGAAGAGTGTCTGAGAA 5773
Db 16626 AGAGGGCCATTTGGCTTGGCTTGAAGTCTTTGGTGATGTGTAGAAGAGTGTCTGAGAA 16567
QY 5774 AGAGAAGGGCCCTGAGCTCGAGGGCAGGCCACCCCTGCGAGTCTGCCAGGCCCTCAG 5833
Db 16566 AGAGAAGGGCCCTGAGCTCGAGGGCAGGCCACCCCTGCGAGTCTGCCAGGCCCTCAG 16507
QY 5834 CCAGCAGTCTGTAGACCCAGGGAGAGACAGGTAGAAGGGCTGGCAGCGAGTGGAGGT 5893
Db 16506 CCAGCAGTCTGTAGACCCAGGGAGAGACAGGTAGAAGGGCTGGCAGCGAGTGGAGGT 16447
QY 5894 GGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCCTGTGGCCCTTCTTGGGTGTGGCCT 5953
Db 16446 GGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCCTGTGGCCCTTCTTGGGTGTGGCCT 16387
QY 5954 GGTGGGCAATTTGTGCGAGCAGGGCAGCTGGGGCTGCTCCCTGAGGATCTCTTGCCTCA 6013
Db 16386 GGTGGGCAATTTGTGCGAGCAGGGCAGCTGGGGCTGCTCCCTGAGGATCTCTTGCCTCA 16327
QY 6014 CCTCATCCAGGGCAATGGCTCAGCCAAAGACCTACTATCGGGATGAACACATCTACCCCT 6073
Db 16326 CCTCATCCAGGGCAATGGCTCAGCCAAAGACCTACTATCGGGATGAACACATCTACCCCT 16267
QY 6074 ACATGTACCTGGCTGGCTACCACTGTGCGAACCCCAATGTGCGGGAAGCCCTGAGGGCT 6133
Db 16266 ACATGTACCTGGCTGGCTACCACTGTGCGAACCCCAATGTGCGGGAAGCCCTGAGGGCT 16207
QY 6134 GGGCGGACAGGGCCACCTGTCTATCCAGGAGTGGAGTCCCTTACTAGGGCTGTCAGGGCTG 6193
Db 16206 GGGCGGACAGGGCCACCTGTCTATCCAGGAGTGGAGTCCCTTACTAGGGCTGTCAGGGCTG 16147
QY 6194 TCTTTCTTCCCTTCCATCAGTTTTCACACACCTCTGTCAGGACTGAGGGCTGGTCCC 6253
Db 16146 TCTTTCTTCCCTTCCATCAGTTTTCACACACCTCTGTCAGGACTGAGGGCTGGTCCC 16087
QY 6254 ACGCCCATCCCTTTCCATCCAGTCCCTAGGAGAGGCCACCATTTACCCAGAGGTA 6313
Db 16086 ACGCCCATCCCTTTCCATCCAGTCCCTAGGAGAGGCCACCATTTACCCAGAGGTA 16027
QY 6314 GGGACCTGATTAAGTGTCTACATCTTCCCTCCCTCCCTCTCTCTCTAAATTTTTTTTT 6373
Db 16026 GGGACCTGATTAAGTGTCTACATCTTCCCTCCCTCCCTCTCTCTCTAAATTTTTTTTT 15967
QY 6374 TCTCAGAACAGTCTCAAAATCTCAATGTTTAAACCAACCATCATCCAGCAGTGGGACTTCCA 6433
Db 15966 TCTCAGAACAGTCTCAAAATCTCAATGTTTAAACCAACCATCATCCAGCAGTGGGACTTCCA 15907
QY 6434 CCCTCGGGCCCATGCCCTCCCTCATCTTGTCTTCTCTGGGTGACCCAGACAG 6493
Db 15906 CCCTCGGGCCCATGCCCTCCCTCATCTTGTCTTCTCTGGGTGACCCAGACAG 15847
QY 6494 CATCATTTTGCAGTGTAGGACCCCACTACTCCCCAGCCCC - TGGGGGCTCCATCCCCCG 6552
Db 15846 CATCATTTTGCAGTGTAGGACCCCACTACTCCCCAGCCCCGCTGGGCTCCATCCCCCG 15787
QY 6553 CCAGGTCCCTGGGCTACCCCGATGGTGAACCCCTTTCAGACCCCTACAGAGACCCCACT 6612
Db 15786 CCAGGTCCCTGGGCTACCCCGATGGTGAACCCCTTTCAGACCCCTACAGAGACCCCACT 15727
QY 6613 GCTCTCACAGCTACAACTACTGCCGGGAAGACGAGAGATCTACAGGAGTCTTTGAAG 6672
Db 15726 GCTCTCACAGCTACAACTACTGCCGGGAAGACGAGAGATCTACAGGAGTCTTTGAAG 15667

QY 6673 TAGCCAAATGATGTCTATCCCAACCTGCTGAAGAGGCGAGCCAGCTGTGTGAGGGGGGCG 6732
Db 15666 TAGCCAAATGATGTCTATCCCAACCTGCTGAAGAGGCGAGCCAGCTGTGTGAGGGGGGCG 15607
QY 6733 AGGAGCGCGGGGAGCAAGCCAGGTGAAGGCTGAGCTCCAGCTGTGTGAGGGGGGCG 6792
Db 15606 AGGAGCGCGGGGAGCAAGCCAGGTGAAGGCTGAGCTCCAGCTGTGTGAGGGGGGCG 15547
QY 6793 CCCACTGGACAGGCTCCCTTCCACAGGGGCGATGGGGCTGCAATGATGAGGATTTAGGGA 6852
Db 15546 CCCACTGGACAGGCTCCCTTCCACA - GGCCATGGGGCTGCAATGATGAGGATTTAGGGA 15488
QY 6853 TGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGTCTATACCTGAGTGC 6912
Db 15487 TGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGTCTATACCTGAGTGC 15428
QY 6913 CTGGCCACAGCTTGGACGGGACTGAAGGTATTTAGAGGTTTCTACCCCTGTGCCCTTCAGT 6972
Db 15427 CTGGCCACAGCTTGGACGGGACTGAAGGTATTTAGAGGTTTCTACCCCTGTGCCCTTCAGT 15368
QY 6973 TTCTATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGTGGATGGTACCTGCTGGCT 7032
Db 15367 TTCTATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGTGGATGGTACCTGCTGGCT 15308
QY 7033 ATGGATGGCTTTTAAAGGAAAGAGGTCTAAGAAATGTTTCCCAACCTATGCTTTACCTT 7092
Db 15307 ATGGATGGCTTTTAAAGGAAAGAGGTCTAAGAAATGTTTCCCAACCTATGCTTTACCTT 15248
QY 7093 TTCTGGAGCCAGGGCTTTTGCCTAGGTGGGGGCTGGCCCTGCTGCTTGAAGGG 7152
Db 15247 TTCTGGAGCCAGGGCTTTTGCCTAGGTGGGGGCTGGCCCTGCTGCTTGAAGGG 15188
QY 7153 TGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTGCTGAGGGCACCCAGAGCAAGG 7212
Db 15187 TGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTGCTGAGGGCACCCAGAGCAAGG 15128
QY 7213 TTCCGCCCTCCAGGACCTGAGTGTCTCCGCCACTGCTGCGATTTCTACGACGGCATCTG 7272
Db 15127 TTCCGCCCTCCAGGACCTGAGTGTCTCCGCCACTGCTGCGATTTCTACGACGGCATCTG 15068
QY 7273 CAAATGGGAGGAGGAGTCCACGCCCTGTGCTGACCTGGGCTGGGCCACTTTCTTGT 7332
Db 15067 CAAATGGGAGGAGGAGTCCACGCCCTGTGCTGACCTGGGCTGGGCCACTTTCTTGT 15008
QY 7333 GCAGTCCCTAGGCCCTTTTGGGAGACAGGTGAGGACAGCTGCACAGAGGTCTGGGCACT 7392
Db 15007 GCAGTCCCTAGGCCCTTTTGGGAGACAGGTGAGGACAGCTGCACAGAGGTCTGGGCACT 14948
QY 7393 ACAGTGGTGACAGCAGCCAGGGCTTGTACAGCTTTTCTGGCCCGGAGGACATCTGC 7452
Db 14947 ACAGTGGTGACAGCAGCCAGGGCTTGTACAGCTTTTCTGGCCCGGAGGACATCTGC 14888
QY 7453 CCATCCCTTCGGTCCGATGGGACTGAGACCCCTGGGTGGGATGGGATGGCCAGAGCA 7512
Db 14887 CCATCCCTTCGGTCCGATGGGACTGAGACCCCTGGGTGGGATGGGATGGCCAGAGCA 14828
QY 7513 GGGTCTTGAGTTCCAGGCACTGGCGGCAACTTGTCTCTACCTTGTCTCTCCCGCACTGG 7572
Db 14827 GGGTCTTGAGTTCCAGGCACTGGCGGCAACTTGTCTCTACCTTGTCTCTCCCGCACTGG 14768
QY 7573 CCCAGGTGGGAGAGGTGGCCATAGTGAAGGCGGAGGCGGCGGCGGAGGCGGCGGAGG 7632
Db 14767 CCCAGGTGGGAGAGGTGGCCATAGTGAAGGCGGAGGCGGCGGCGGAGGCGGCGGAGG 14708
QY 7633 AGCCGTGGGCGAGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGC 7692
Db 14707 AGCCGTGGGCGAGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGC 14648
QY 7693 CAGAGGAGCCCGCGGCCCAAGAGCCAGCAGCTGGACAAGGGCTTGGGCGACCGGCCAGG 7752
Db 14647 CAGAGGAGCCCGCGGCCCAAGAGCCAGCAGCTGGACAAGGGCTTGGGCGACCGGCCAGG 14588
QY 7753 GTGCAGTGTACAGACCCCGGGAAGCCCTCTTGGGACTGTGCTGGCAGACAGCCCGGAGGCC 7812

Db 14587 GTGCAGTGTACAGACCCCGGAGAGCCCTCTGTGGACTGTGCTGGACAGCCCGAGGCC 14528
QY 7813 CTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCGCAGCATCACACCGCGGAGGCTC 7872
Db 14527 CTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCGCAGCATCACACCGCGGAGGCTC 14468
QY 7873 CAGTGTCTCAGCTTTCCAGAGTGAAGATGAAGGGCATGAAGGGAGCTGTGTGGGCCACCA 7932
Db 14467 CAGTGTCTCAGCTTTCCAGAGTGAAGATGAAGGGCATGAAGGGAGCTGTGTGGGCCACCA 14408
QY 7933 AGATCAACTCAGCGCCCATCAAGCTGCAACTCAGCGCACAGTGCGAAGTGCAGATGAAGA 7992
Db 14407 AGATCAACTCAGCGCCCATCAAGCTGCAACTCAGCGCACAGTGCGAAGTGCAGATGAAGA 14348
QY 7993 AGCAGAAAGTGTCCACCCCTAGTGACTACACTCTCTCTTCTCAAGCGGAGCCCAAG 8052
Db 14347 AGCAGAAAGTGTCCACCCCTAGTGACTACACTCTCTCTTCTCAAGCGGAGCCCAAG 14288
QY 8053 GCCTCTGAACCTACTGGGACTTCCGAGCCTTGTGGGACCCAGGCTCCG - CCTTTAGTCC 8111
Db 14287 GCCTCTGAACCTACTGGGACTTCCGAGCCTTGTGGGACCCAGGCTCCGCTTTAGTCC 14228
QY 8112 CCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGAGAGCCCTCACCTCTACCCA 8171
Db 14227 CCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGAGAGCCCTCACCTCTACCCA 14168
QY 8172 AACCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTCTCCAGCTCA 8231
Db 14167 AACCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTCTCCAGCTCA 14108
QY 8232 GAACCCAGGGCTCTGCCCGAGTCTAGATATAGGTCTCTTCTCCAGAAATCCAGCGG 8291
Db 14107 GAACCCAGGGCTCTGCCCGAGTCTAGATATAGGTCTCTTCTCCAGAAATCCAGCGG 14048
QY 8292 GCCAATGGAACCTCACGCTGGGTCTTAATACAGTCTTTAAAGGCCAGGCCCTAGAA 8351
Db 14047 GCCAATGGAACCTCACGCTGGGTCTTAATACAGTCTTTAAAGGCCAGGCCCTAGAA 13988
QY 8352 ACCAAGCTCTCTCGGAGCGCTCACCTAGAGCCAGCAACCGTTACTCAGGSGCTCT 8411
Db 13987 ACCAAGCTCTCTCGGAGCGCTCACCTAGAGCCAGCAACCGTTACTCAGGSGCTCT 13928
QY 8412 CCCAGCTTGTAGGAGCTGAGGTTTTCACCTTTAACCCAAAGGAGCAGGTCCCAGCTCCA 8471
Db 13927 CCCAGCTTGTAGGAGCTGAGGTTTTCACCTTTAACCCAA - GGAGCACAGGTCCCAGCTCCA 13869
QY 8472 GCCGGGAGGCTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACCTTCAGAATTTCA 8531
Db 13868 GCCC - GGAGGCTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACCTTCAGAATTTCA 13810
QY 8532 TATCTTGGCAATCCAAAGCTCCCTGCCCAATTAACCTTCAGTCTCTCTTCCAGATTTTGA 8591
Db 13809 TATCTTGGCAATCCAAAGCTCCCTGCCCAATTAACCTTCAGTCTCTCTTCCAGATTTTGA 13750
QY 8592 AATCTAGTCTTCTCTCTCCTC 8612
Db 13749 TCCTTCTCTCTCTCTCTCTC 13729

RESULT 6
AP001187/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-665N17 map 11q13, WORKING
DRAFT SEQUENCE, 27 unordered pieces.
ACCESSION AP001187
VERSION AP001187.3 GI:12246852
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone: RP11-665N17.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 166269)

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 166,269 genomic DNA of 11q13
 Published Only in DataBase (2000) In press
 2 (bases 1 to 166269)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (18-FEB-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel:81-45-503-9111, Fax:81-45-503-9170
 E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 On Jan 16, 2001 this sequence version replaced gi:8117633.
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsr.riken.go.jp
 ----- Project Information
 Center project name: HumDraft11
 Center clone name: RP11-665N17
 ----- Summary Statistics

COMMENT

Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amerham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 156782 bases at least Q40
 Consensus quality: 160361 bases at least Q30
 Consensus quality: 162264 bases at least Q20
 Insert size: 163669; sum-of-contigs
 Quality coverage: 8.32x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 27 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

```

1      28395 contig of 28395 bp in length
      28496      43446 contig of 14951 bp in length
      43547      58083 contig of 14537 bp in length
      58184      66868 contig of 8685 bp in length
      66969      76574 contig of 9606 bp in length
      76675      86418 contig of 9744 bp in length
      86519      94044 contig of 7526 bp in length
      94145      100790 contig of 6646 bp in length
      100891      108540 contig of 7649 bp in length
      108640      115961 contig of 100 bp in length
      115962      120258 contig of 4114 bp in length
      120259      125359 contig of 5001 bp in length
      125360      130186 contig of 4727 bp in length
      130187      134400 contig of 1044 bp in length
      134401      139064 contig of 1174 bp in length
      139065      143906 contig of 1048 bp in length
      143907      148280 contig of 1543 bp in length
      148281      151006 contig of 2626 bp in length
      151007      155555 contig of 1707 bp in length
      155556      157198 contig of 1543 bp in length
      157199      159872 contig of 2574 bp in length
      159873      161499 contig of 1527 bp in length
      161500      162703 contig of 1104 bp in length
      162704      163847 contig of 1044 bp in length
      163848      165121 contig of 1044 bp in length
      165122      166269 contig of 1048 bp in length.
  
```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

```

* be preserved.
*      1      28395: contig of 28395 bp in length
*      28396      28495: gap of 100 bp
*      28496      43446: contig of 14951 bp in length
*      43447      43546: gap of 100 bp
*      43547      58083: contig of 14537 bp in length
*      58084      58183: gap of 100 bp
*      58184      66868: contig of 8685 bp in length
*      66869      66968: gap of 100 bp
*      66969      76574: contig of 9606 bp in length
*      76575      76674: gap of 100 bp
*      76675      86418: contig of 9744 bp in length
*      86419      86518: gap of 100 bp
*      86519      94044: contig of 7526 bp in length
*      94045      94144: gap of 100 bp
*      94145      100790: contig of 6646 bp in length
*      100791      100890: gap of 100 bp
*      100891      108539: contig of 7649 bp in length
*      108540      108639: gap of 100 bp
*      108640      115961: contig of 7322 bp in length
*      115962      116061: gap of 100 bp
*      116062      120258: contig of 4197 bp in length
*      120259      120358: gap of 100 bp
*      120359      125359: contig of 5001 bp in length
*      125360      125459: gap of 100 bp
*      125460      130186: contig of 4727 bp in length
*      130187      130286: gap of 100 bp
*      130287      134400: contig of 4114 bp in length
*      134401      134500: gap of 100 bp
*      134501      139064: contig of 4564 bp in length
*      139065      139164: gap of 100 bp
*      139165      143906: contig of 4742 bp in length
*      143907      144006: gap of 100 bp
*      144007      148280: contig of 4274 bp in length
*      148281      148380: gap of 100 bp
*      148381      151006: contig of 2626 bp in length
*      151007      151106: gap of 100 bp
*      151107      153748: contig of 2642 bp in length
*      153749      153848: gap of 100 bp
*      153849      155555: contig of 1707 bp in length
*      155556      155855: gap of 100 bp
*      155856      157198: contig of 1543 bp in length
*      157199      157298: gap of 100 bp
*      157299      159872: contig of 2574 bp in length
*      159873      159972: gap of 100 bp
*      159973      161499: contig of 1527 bp in length
*      161500      161599: gap of 100 bp
*      161600      162703: contig of 1104 bp in length
*      162704      162803: gap of 100 bp
*      162804      163847: contig of 1044 bp in length
*      163848      163947: gap of 100 bp
*      163948      165121: contig of 1174 bp in length
*      165122      165221: gap of 100 bp
*      165222      166269: contig of 1048 bp in length.
  
```

FEATURES

```

source
  1..166269
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q13"
    /clone="RP11-665N17"
  misc_feature
    1..28395
      /note="assembly_fragment"
  misc_feature
    28496..43446
      /note="assembly_fragment"
  misc_feature
    43547..58083
      /note="assembly_fragment"
  misc_feature
    58184..66868
      /note="assembly_fragment clone_end:T7 vector_side:right"
  misc_feature
    66969..76574
      /note="assembly_fragment"
  misc_feature
    76675..86418
      /note="assembly_fragment"
  
```

[illegible]

Db 137738 CCAGACAGTCAATGCCGGTGTGGCTGAGCGGTAATTTGTTCCCTCCCCACGCTTGTCC 137679
QY 4535 CTTTACTACTGTAGTACCCAAAGCCACCCAAAGGACTCCATTTTCTTTGGGCGACACCCCTT 4594
Db 137678 CTTTACTACTGTAGTACCCAAAGCCACCCAAAGGACTCCATTTTCTTTGGGCGACACCCCTT 137619
QY 4595 TCTTCCCATCACACCACATAGGAAGGAAAGACAGAAGAGCCCCCTTTTCTTGGCTGTCA 4654
Db 137618 TCTTCCCATCACACCACATAGGAAGGAAAGACAGAAGAGCCCCCTTTTCTTGGCTGTCA 137559
QY 4655 TTCCCTGAAGCAGGCACAGGGTGGGCCATCATGACAGACATAAATCATCTCATCCCCCCTAA 4714
Db 137558 TTCCCTGAAGCAGGCACAGGGTGGGCCATCATGACAGACATAAATCATCTCATCCCCCCTAA 137499
QY 4715 GAGCTGGCTGTACTCTAAAGAGTATACATATCGCTGTGACCGCAAGATGGAGGTGGCGTT 4774
Db 137498 GAGCTGGCTGTACTCTAAAGAGTATACATATCGCTGTGACCGCAAGATGGAGGTGGCGTT 137439
QY 4775 CATGGTGTGGCCATCAACCCCTTCCATTTGACCTTGCACACCGACTCGCTGGAGCTTCTGCA 4834
Db 137438 CATGGTGTGGCCATCAACCCCTTCCATTTGACCTTGCACACCGACTCGCTGGAGCTTCTGCA 137379
QY 4835 GCTGCAGCAGGTGAGGGCTGAGCCATGGGCAGGACTGGGCTAGGCCAGACTTGACTTG 4894
Db 137378 GCTGCAGCAGGTGAGGGCTGAGCCATGGGCAGGACTGGGCTAGGCCAGACTTGACTTG 137319
QY 4895 CTGTGGACCCCTGGCAGGGGCACTTTTCCCTTCTGAGCTTTCAGCTTCCCTCGCTGGA 4954
Db 137318 CTGTGGACCCCTGGCAGGGGCACTTTTCCCTTCTGAGCTTTCAGCTTCCCTCGCTGGA 137259
QY 4955 AATGGGTTAGTAATTCCTGGCTGGCTTTCCAGGGCTTTCCAGGGCTTTGGGAGAGTAGAATTGAGA 5014
Db 137258 AATGGGTTAGTAATTCCTGGCTGGCTTTCCAGGGCTTTCCAGGGCTTTGGGAGAGTAGAATTGAGA 137199
QY 5015 TGTGAATTTGCTTTGACTCCATTAAGGGCTGTGCCAGAAATTTTGGCCCTTCCACATGG 5074
Db 137198 TGTGAATTTGCTTTGACTCCATTAAGGGCTGTGCCAGAAATTTTGGCCCTTCCACATGG 137139
QY 5075 TGGGTGGTCCCTGCTGTTGTCTGACCCCACTCTGCGCGATAGGCTTAAGGACCCGTTCTC 5134
Db 137138 TGGGTGGTCCCTGCTGTTGTCTGACCCCACTCTGCGCGATAGGCTTAAGGACCCGTTCTC 137079
QY 5135 CTCCTCTGTTCCGTTGCTCATACTCTCTCTGCGCTCTCTAGAGCTGCTTGGCTGCTC 5194
Db 137078 CTCCTCTGTTCCGTTGCTCATACTCTCTCTGCGCTCTCTAGAGCTGCTTGGCTGCTC 137019
QY 5195 TATGACCTGGGACATCTGGAAGGTCAGTAGAGGGAAGTGGCCAGGCTGGCCCTGGTGAG 5254
Db 137018 TATGACCTGGGACATCTGGAAGGTCAGTAGAGGGAAGTGGCCAGGCTGGCCCTGGTGAG 136959
QY 5255 GCCGGGGGCTGGTGGCAGCCTGAATATGATCTTCTCTAGTACCCATGCCCTTAG 5314
Db 136958 GCCGGGGGCTGGTGGCAGCCTGAATATGATCTTCTCTAGTACCCATGCCCTTAG 136899
QY 5315 GGAACCTGGCAGATCTAGAGGACTGGAGCCCAACCCCTGGCCGGCCAGACCCACTCACCC 5374
Db 136898 GGAACCTGGCAGATCTAGAGGACTGGAGCCCAACCCCTGGCCGGCCAGACCCACTCACCC 136839
QY 5375 TCTACACAAGGTGGGGGATCTAAGAGGTCAGAGGTCAGAGGGAGACCCCTAACAGTGGCTGA 5434
Db 136838 TCTACACAAGGTGGGGGATCTAAGAGGTCAGAGGTCAGAGGGAGACCCCTAACAGTGGCTGA 136779
QY 5435 GGCAGGGGCCCTCATCTGGCAGATCAGAGAGAACTTTGTGTGGGGGTATCGCC 5494
Db 136778 GGCAGGGGCCCTCATCTGGCAGATCAGAGAGAACTTTGTGTGGGGGTATCGCC 136719
QY 5495 ATCCAGTCTCAGTTTGTGTCAACTGTGTGAGAAATCAGTTTCAAGGCTGTCTGAGGG 5554
Db 136718 ATCCAGTCTCAGTTTGTGTCAACTGTGTGAGAAATCAGTTTCAAGGCTGTCTGAGGG 136659
QY 5555 GTCTCAGGGTTCCCGAGCTGGAGTGGCAGGGGTGCAATTTGTCCCTTCAGCCCTGCC 5614
Db 136658 GTCTCAGGGTTCCCGAGCTGGAGTGGCAGGGGTGCAATTTGTCCCTTCAGCCCTGCC 136599

QY 5615 TTTTCTGCCACTGCTTACTGTCTCTCTGAGTATACAGAGTCAAAATCTGCTAGGAGC 5674
Db 136598 TTTTCTGCCACTGCTTACTGTCTCTCTGAGTATACAGAGTCAAAATCTGCTAGGAGC 136539
QY 5675 AC---TCAAGAGGGGTGTTCACCTTGGTGGGTGTAGTGGGAGGAGGCCATTTGGGCTG 5731
Db 136538 ACTGATGAAGAGGGGTGTTCACCTTGGTGGGTGTAGTGGGAGGAGGCCATTTGGGCTG 136479
QY 5732 GGCTTGAAGTCTTTGGTGTATGTGTAGAACAGTGTCTGAGAAAGAGAGGCCCTGAGCT 5791
Db 136478 GGCTTGAAGTCTTTGGTGTATGTGTAGAACAGTGTCTGAGAAAGAGAGGCCCTGAGCT 136419
QY 5792 CGAGGCGAGGCCCCACCCCTGCAGTCTGCCAGGCTCAGCCAGCAGTCTCTGTAGACC 5851
Db 136418 CGAGGCGAGGCCCCACCCCTGCAGTCTGCCAGGCTCAGCCAGCAGTCTCTGTAGACC 136359
QY 5852 CAGGGAGGAGACCAAGGTAGAGGGCTGGCAGCGAGTGGAGTGGGAGTGGAGATGGAGAG 5911
Db 136358 CAGGGAGGAGACCAAGGTAGAGGGCTGGCAGCGAGTGGAGTGGGAGTGGAGAG 136299
QY 5912 GACTCCCTGGGATCTTCCCTGTGGCCCTTCTGTGGTGTGCCCTGTGGGCAATTTGTGCCA 5971
Db 136298 GACTCCCTGGGATCTTCCCTGTGGCCCTTCTGTGGTGTGCCCTGTGGGCAATTTGTGCCA 136239
QY 5972 GCAGGCGAGCTGGGGCTGCCCTCTGAGGATCTCTGCCCTCAGCTTCCATCCAGGCAATTG 6031
Db 136238 GCAGGCGAGCTGGGGCTGCCCTCTGAGGATCTCTGCCCTCAGCTTCCATCCAGGCAATTG 136179
QY 6032 CTTACGCAAGACCTACTATFCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCT 6091
Db 136178 CTTACGCAAGACCTACTATFCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCT 136119
QY 6092 ACCACTGTCCAAACCGCAATGTGGGGAGCCCTGACAGGCTGGGGCGACACGCCCACTG 6151
Db 136118 ACCACTGTCCAAACCGCAATGTGGGGAGCCCTGACAGGCTGGGGCGACACGCCCACTG 136059
QY 6152 TCATCAGAGAGTGAAGTATCCCTACTAGGCGCTGACAGCCTGTCTTCTTCTCCCTCCAT 6211
Db 136058 TCATCAGAGAGTGAAGTATCCCTACTAGGCGCTGACAGCCTGTCTTCTTCTCCCTCCAT 135999
QY 6212 CAGTTTCCAAACCACTCTGCTCAGGACTGAGGCTGGCTCCCAAGCCCACTCCCTTTCC 6271
Db 135998 CAGTTTCCAAACCACTCTGCTCAGGACTGAGGCTGGCTCCCAAGCCCACTCCCTTTCC 135939
QY 6272 ATCCAGTCCCTAGGCAAGGCCACCATTAACCCAGAGGTAGGGACCTTGATTAAGGTG 6331
Db 135938 ATCCAGTCCCTAGGCAAGGCCACCATTAACCCAGAGGTAGGGACCTTGATTAAGGTG 135879
QY 6332 TCACATCTTTCCCTCCCTCCCTCTCTCTCTAAATTTTCTCTAGAACAGTCTCAAA 6391
Db 135878 TCACATCTTTCCCTCCCTCCCTCTCTCTCTAAATTTTCTCTAGAACAGTCTCAAA 135819
QY 6392 TCTCCAAATGTTTAAACCACTATCCAGCAGTGGGACTTCCACCTCGGCCCACTGCC 6451
Db 135818 TCTCCAAATGTTTAAACCACTATCCAGCAGTGGGACTTCCACCTCGGCCCACTGCC 135759
QY 6452 CTTCTCATTTCTTCTTCTCTCTGGGCTGACCCAGACAGATCATTTTGCAGTGG 6511
Db 135758 CTTCTCATTTCTTCTTCTCTCTGGGCTGACCCAGACAGATCATTTTGCAGTGG 135699
QY 6512 ACCCCACTACTCCCCAGCCCTGGGGCTTCCATCCCCCGCAGGTCCTTGGGGCTACC 6571
Db 135698 ACCCCACTACTCCCCAGCCCTGGGGCTTCCATCCCCCGCAGGTCCTTGGGGCTACC 135639
QY 6572 CCCGATGGTGAAGCCCTTCAGACCCCTTACAGAGACCCCACTGCTTCCACAGTACAATA 6631
Db 135638 CCCGATGGTGAAGCCCTTCAGACCCCTTACAGAGACCCCACTGCTTCCACAGTACAATA 135579
QY 6632 CTGCCGGGAAGACGAGGAGATCTAAGAGGATCTTTGAAGTAGCCCAATGATCATCC 6691
Db 135578 CTGCCGGGAAGACGAGGAGATCTAAGAGGATCTTTGAAGTAGCCCAATGATCATCC 135519

NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	20516:	contig of 20516 bp in length
20517	20616:	gap of 100 bp
20617	30813:	contig of 10197 bp in length
30814	30913:	gap of 100 bp
30914	42329:	contig of 11416 bp in length
42420	42429:	gap of 100 bp
42430	53053:	contig of 10624 bp in length
53054	53153:	gap of 100 bp
53154	63766:	contig of 10613 bp in length
63767	73866:	gap of 100 bp
73867	73322:	contig of 9456 bp in length
73423	73422:	gap of 100 bp
73423	81331:	contig of 7909 bp in length
81332	81431:	gap of 100 bp
81432	88430:	contig of 6999 bp in length
88431	88530:	gap of 100 bp
88531	93403:	contig of 4873 bp in length
93404	93503:	gap of 100 bp
93504	98277:	contig of 4774 bp in length
98278	98377:	gap of 100 bp
98378	101730:	contig of 3353 bp in length
101731	101830:	gap of 100 bp
106311	106170:	contig of 4340 bp in length
106312	106270:	gap of 100 bp
106271	110043:	contig of 3773 bp in length
110144	110143:	gap of 100 bp
110144	1146327:	contig of 4484 bp in length
114628	114727:	gap of 100 bp
114728	117498:	contig of 2771 bp in length
117499	117598:	gap of 100 bp
117599	120318:	contig of 2720 bp in length
120319	120418:	gap of 100 bp
120419	122917:	contig of 2499 bp in length
122918	123017:	gap of 100 bp
123018	125107:	contig of 2090 bp in length
125108	125207:	gap of 100 bp
125208	127881:	contig of 2674 bp in length
127882	127981:	gap of 100 bp
127982	130814:	contig of 2833 bp in length
130815	130914:	gap of 100 bp
130915	132593:	contig of 1679 bp in length
132594	132693:	gap of 100 bp
132894	134516:	contig of 1823 bp in length
134517	134616:	gap of 100 bp
134617	137106:	contig of 2490 bp in length
137107	137206:	gap of 100 bp
137207	140498:	contig of 3292 bp in length
140499	140598:	gap of 100 bp
140599	143489:	contig of 2891 bp in length
143490	143589:	gap of 100 bp
143590	145559:	contig of 1970 bp in length
145560	145659:	gap of 100 bp
145660	147741:	contig of 2082 bp in length
147742	147841:	gap of 100 bp
147842	149318:	contig of 1477 bp in length
149319	149418:	gap of 100 bp
149419	150896:	contig of 1478 bp in length
150897	150996:	gap of 100 bp
150997	152442:	contig of 1446 bp in length
152443	152542:	gap of 100 bp
152543	154192:	contig of 1650 bp in length
154292	154292:	gap of 100 bp
154293	156330:	contig of 2038 bp in length

40070 CC9CAC190100CC1A19AAAA1111111111

b 48870 CCGCAACTGGTGGCCTATGAAAATTTTTTTTTTTTTCAGACGGCGCTCACTCTGTCG 48929

[illegible]

Qy	6743	GGGGAGCAAAAGCCAGGTGAAAGGCTGGAGCTCCAGGCTGTGTCCAGCCTCCCACTGGGA	6802
Db	52226	GGGGAGCAAAAGCCAGGTGAAAGGCTGGAGCTCCAGGCTGTGTCCAGCCTCCCACTGGGA	52285
Qy	6803	CAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTACGGGATAGGGATGGCAGGAGG	6862
Db	52286	CAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTACGGGATAGGGATGGCAGGAGG	52344
Qy	6863	AAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTGGCTATAACTCAGGTCTCTGGGCCAC	6922
Db	52345	AAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTGGCTATAACTCAGGTCTCTGGGCCAC	52404
Qy	6923	GTTGACGGGACTGAAAGTATTTTAGAGGTTTCTACCCCTGTGCCTTTCAGTTTCATGGCCA	6982
Db	52405	GTTGACGGGACTGAAAGTATTTTAGAGGTTTCTACCCCTGTGCCTTTCAGTTTCATGGCCA	52464
Qy	6983	GACTCCCTCCCTCAGCTGAGGGGTGGAGGTAGGATGGTACGTCTGCTGCTATGGATTGGC	7042
Db	52465	GACTCCCTCCCTCAGCTGAGGGGTGGAGGTAGGATGGTACGTCTGCTGCTATGGATTGGC	52524
Qy	7043	TTTATAAAGAAAGAGAGGTTCTAAGAATGTTCCCAACCTATGCTTTTCTTCTGGAGCC	7102
Db	52525	TTTATAAAGAAAGAGAGATTTAAGAATGTTCCCAACCTATGCTTTTCTTCTGGAGCC	52584
Qy	7103	AGGGGCTTTTGCTTAGGTGGGGGCTGGCCCTGTGCCCTCTGCTAAGGGGTGAGTAAGAG	7162
Db	52585	AGGGGCTTTTGCTTAGGTGGGGGCTGGCCCTGTGCCCTCTGCTAAGGGGTGAGTAAGAG	52644
Qy	7163	ACTGATCTGTGCCCTCCCTTCCCTCTGCTCAGGGGACCCAGAGCAAGTTTCGCGCCCTC	7222
Db	52645	ACTGATCTGTGCCCTCCCTTCCCTCTGCTCAGGGGACCCAGAGCAAGTTTCGCGCCCTC	52704
Qy	7223	CAGGACCTGAGTGTCTGCGCCCACTGCTGCGATTTCTACGACGGCACTCTGCAAAATGGGAG	7282
Db	52705	CAGGACCTGAGTGTCTGCGCCCACTGCTGCGATTTCTACGACGGCACTCTGCAAAATGGGAG	52764
Qy	7283	GAGGCGAGTCCCAAGCCCTGTGCTGCACGTGGGCTGGGCCACCTTTCTTTGTGCAGTCCCTA	7342
Db	52765	GAGGCGAGTCCCAAGCCCTGTGCTGCATGTGGCTGGGCCACCTTTCTTTGTGCAGTCCCTA	52824
Qy	7343	GGCCGTTTGTAGGACAGGTGAGGACAGCTGACAGAGGTCTGGGCACTACAGGTGTGTG	7402
Db	52825	GGCCGTTTGTAGGACAGGTGAGGACAGCTGACAGAGGTCTGGGCACTACAGGTGTGTG	52884
Qy	7403	ACAGCAGCCAGGGCTGCTCAGACTTTTCTGGCCAGGGGAGCATCTGCCCATCCCTT	7462
Db	52885	ACAGCAGCCAGGGCTGCTCAGACTTTTCTGGCCAGGGGAGCATCTGCCCATCCCTT	52944
Qy	7463	CGGTGCCGATGGGACTGAGACCCCTGGGTGGGATGGGATGCCAGACAGGTCCTTGGGA	7522
Db	52945	GGGTGCCGATGGGACTGAGACCCCTGGGTGGGATGGGATGCCAGACAGGTCCTTGGGA	53004
Qy	7523	GTTCCAGCCACTGGCCGCAACCTTTGCTCTCACCCTTGTCTCTCCCACTGGGCCAGGTCGG	7582
Db	53005	GTTCCAGCCACTGGCCGCAACCTTTGCTCTCACCCTTGTCTCTCCCACTGGGCCAGGTCGG	53064
Qy	7583	GCAGAAGTGCGCATAGTGTAGCCGAGAGGCGCGCCGAGGCGCGGAGGCGGTGGG	7642
Db	53065	NN	53124
Qy	7643	CGAGGAAGCCCGGAAGGCGCGCGGGGCC 7675	
Db	53125	NN	53157
RESULT	8		
LOCUS	AP001558		
DEFINITION	AP001558	178100 bp	DNA
ACCESSION	AP001558		HTG
VERSION	AP001558.2		30-MAY-2000
			map 11q13, WORKING
			DRAFT SEQUENCE, 40 unordered pieces.

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Homo sapiens DNA, clone:RP11-686D5.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178100)

REFERENCE

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,

AUTHORS

Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Homo sapiens 178,100 genomic DNA of 11q13

REFERENCE

Published Only in Database (2000) In press

AUTHORS

2 (bases 1 to 178100)

TITLE

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,

JOURNAL

Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-2000) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

Kitasato Univ. 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,

Japan (E-mail: hattori@gsc.riken.go.jp,

URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,

Fax: 81-42-778-9924)

COMMENT

On May 30, 2000 this sequence version replaced gi:7380893.

Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: Humdraft11

Center clone name: RP11-686D5

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 148042 bases at least Q40

Consensus quality: 162125 bases at least Q30

Consensus quality: 169512 bases at least Q20

Insert size: 174200; sum-of-contigs

Quality coverage: 4.12x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 40 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

157919 159802 contig of 1884 bp in length
159903 161284 contig of 1382 bp in length
161385 163110 contig of 1726 bp in length
164636 contig of 1426 bp in length
166593 contig of 1857 bp in length
167867 contig of 1174 bp in length
169543 contig of 1576 bp in length
170824 contig of 1181 bp in length
172925 172804 contig of 1880 bp in length
172905 173997 contig of 1093 bp in length
174098 175750 contig of 1653 bp in length
175851 176984 contig of 1134 bp in length
177085 178100 contig of 1016 bp in length
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 20070: contig of 20070 bp in length
* 20071 20170: gap of 100 bp
* 20171 34749: contig of 14579 bp in length
* 34750 34849: gap of 100 bp
* 34850 50049: contig of 15200 bp in length
* 50050 50149: gap of 100 bp
* 50150 60765: contig of 10616 bp in length
* 60766 60865: gap of 100 bp
* 60866 67727: contig of 6862 bp in length
* 67728 67827: gap of 100 bp
* 67828 76362: contig of 8535 bp in length
* 76363 76462: gap of 100 bp
* 76463 82230: contig of 5768 bp in length
* 82231 82330: gap of 100 bp
* 82331 89521: contig of 7191 bp in length
* 89522 89621: gap of 100 bp
* 89622 95129: contig of 5508 bp in length
* 95130 95229: gap of 100 bp
* 95230 101095: contig of 5866 bp in length
* 101096 101195: gap of 100 bp
* 101196 106255: contig of 5060 bp in length
* 106256 106355: gap of 100 bp
* 106356 110801: contig of 4446 bp in length
* 110802 110901: gap of 100 bp
* 110902 114837: contig of 3936 bp in length
* 114838 114937: gap of 100 bp
* 114938 120430: contig of 5493 bp in length
* 120431 120530: gap of 100 bp
* 120531 125238: contig of 4708 bp in length
* 125239 125338: gap of 100 bp
* 125339 128370: contig of 3032 bp in length
* 128371 128470: gap of 100 bp
* 128471 131184: contig of 2714 bp in length
* 131185 131284: gap of 100 bp
* 131285 134627: contig of 3343 bp in length
* 134628 134727: gap of 100 bp
* 134728 137446: contig of 2719 bp in length
* 137447 137546: gap of 100 bp
* 137547 140281: contig of 2735 bp in length
* 140282 140381: gap of 100 bp
* 140382 142601: contig of 2220 bp in length
* 142602 142701: gap of 100 bp
* 142702 145477: contig of 2776 bp in length
* 145478 145577: gap of 100 bp
* 145578 148565: contig of 2988 bp in length
* 148566 148665: gap of 100 bp
* 148666 151182: contig of 2517 bp in length
* 151183 151282: gap of 100 bp
* 151283 153836: contig of 2554 bp in length
* 153837 153936: gap of 100 bp
* 153937 156236: contig of 2300 bp in length
* 156237 156337: contig of 2300 bp in length

Qy 4803 GACCTGCACACCGACTCGCTGGAGCTTCTGACGTGACAGAGGTGAGGGCTGAGCCCAATG 4862
|||||
Db 47918 GACCTGCACACCGACTCGCTGGAGCTTCTGACGTGACAGAGGTGAGGGCTGAGCCCAATG 47977
Qy 4863 GGGCAGAGACTGGGCTAGGCCAGACTTGACTTGCTGCTGGGACCCCTGGGCAGGGGCACTTTC 4922
|||||
Db 47978 GGGCAGAGACTGGGCTAGGCCAGACTTGACTTGCTGCTGGGACCCCTGGGCAGGGGCACTTTC 48037
Qy 4923 CTTTCTGAGCTTCAGCTTCCCTCCCTCGGAAATAATGGTTAGTAATTCCTGGCCCTGGCCT 4982
|||||
Db 48038 CTTTCTGAGCTTCAGCTTCCCTCCCTCGGAAATAATGGTTAGTAATTCCTGGCCCTGGCCT 48097
Qy 4983 TTCCAGGGCTCTTGGAGAGTAGAATTAGAGATGAAATTTGCTTTGACTCCATTAAAGG 5042
|||||
Db 48098 TTCCAGGGCTCTTGGAGAGTAGAATTAGAGATGAAATTTGCTTTGACTCCATTAAAGG 48157
Qy 5043 GCTGGTCCAGAAATTTGGCCCTTCCACATGGTGGGTGGTCCCTGTTGGTTCTGACCCCC 5102
|||||
Db 48158 GCTGGTCCAGAAATTTGGCCCTTCCACATGGTGGGTGGTCCCTGTTGGTTCTGACCCCC 48217
Qy 5103 ACCTCTGCCGATAGGCTAAGGACCCGTTCTCCCTCGTTCCTGCTGGCTCATTAACCTCTCT 5162
|||||
Db 48218 ACCTCTGCCGATAGGCTAAGGACCCGTTCTCCCTCGTTCCTGCTGGCTCATTAACCTCTCT 48277
Qy 5163 CTTTCCGCTCCTAGAACTGCTCTGGCTGCTCTATGACCTGGGACATCTGGAAAGGTGAG 5222
|||||
Db 48278 CTTTCCGCTCCTAGAACTGCTCTGGCTGCTCTATGACCTGGGACATCTGGAAAGGTGAG 48337
Qy 5223 TAGAGGGAAGTGGCCAGGCTCGCTGCTGAGGCGCGGGGCTGGGTGGAGCCCTGAATTT 5282
|||||
Db 48338 TAGAGGGAAGTGGCCAGGCTCGCTGCTGAGGCGCGGGGCTGGGTGGAGCCCTGAATTT 48397
Qy 5283 ATGATCTCTTCTTAGTACCCCATGGCTTAGGGAACCTGGCAGATCTAGAGGAGCTGGA 5342
|||||
Db 48398 ATGATCTCTTCTTAGTACCCCATGGCTTAGGGAACCTGGCAGATCTAGAGGAGCTGGA 48457
Qy 5343 GCCACCCCTGGCGGCGAGACCACATCACCTCTACACAAAGGTGGGGCATCTAAGGA 5402
|||||
Db 48458 GCCACCCCTGGCGGCGAGACCACATCACCTCTACACAAAGGTGGGGCATCTAAGGA 48517
Qy 5403 GGGTGCAGAGGGAGCCCTAACAGTGGCTGAGGCGAGGGCCCTCATCTGGGCGAGATGAG 5462
|||||
Db 48518 GGGTGCAGAGGGAGCCCTAACAGTGGCTGAGGCGAGGGCCCTCATCTGGGCGAGATGAG 48577
Qy 5463 AAGAGAACTTTGTGTTGGGGGTATCGCCCATCCAGTCTCACTTTGTGTCAACTGTGT 5522
|||||
Db 48578 AAGAGAACTTTGTGTTGGGGGTATCGCCCATCCAGTCTCACTTTGTGTCAACTGTGT 48637
Qy 5523 GCAGAACTCAGTTCACTCAGGCTGTCTGAGGGGTGTCCAGGGTTCCCGAGCCCTGGGAGTG 5582
|||||
Db 48638 GCAGAACTCAGTTCACTCAGGCTGTCTGAGGGGTGTCCAGGGTTCCCGAGCCCTGGGAGTG 48697
Qy 5583 GCAGGGCTGCAATTTGCTCCCTCAGCCCTGCTTTTCTGCCCACCTGCTTACTGCTCTTCCCT 5642
|||||
Db 48698 GCAGGGCTGCAATTTGCTCCCTCAGCCCTGCTTTTCTGCCCACCTGCTTACTGCTCTTCCCT 48757
Qy 5643 GGAGTATAACAGAGTCAAAATGTTGGTAGGAGCAC - - - TGAAGAGGGGGTGTTCACCTTGGT 5699
|||||
Db 48758 GGAGTATAACAGAGTCAAAATGTTGGTAGGAGCACTGATGAAGAGGGGTGTTTCACTTGGT 48817
Qy 5700 GGGTGTAGGTGGGAGGAGGCCATTTGGCTGGGCTTGAAGTCTTTGGTATGTGTAGA 5759
|||||
Db 48818 GGGTGTAGGTGGGAGGAGGCCATTTGGGCTGGGCTTGAAGTCTTTGGTATGTGTAGA 48877
Qy 5760 AGAGTGTCTGAGAAGAGAAGGGCCCTGAGCTCGAGGGGAGGCCACCCCTCGAGTCT 5819
|||||
Db 48878 AGAGTGTCTGAGAAGAGAAGGGCCCTGAGCTCGAGGGGAGGCCACCCCTCGAGTCT 48937
Qy 5820 GCCCAGGCCCTCAGCCAGCTGCTGTAGACCCAGGGAGGAGACCAAGGTAGAAAGGGCTGG 5879
|||||
Db 48938 GCCCAGGCCCTCAGCCAGCTGCTGTAGACCCAGGGAGGAGACCAAGGTAGAAAGGGCTGG 48997
Qy 5880 CAGCGAGTGGAGGTGGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCTCTGGCCCT 5939

Db 48998 CAGCGAGTGGAGGTGGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCTCTGTGGCCCT 49057
Qy 5940 TCTGGGTGTGCCCTGGTGGGCAATTTGTGCCAGAGGCACTGGGCTGCTCTCCCTGAG 5999
Db 49058 TCTGGGTGTGCCCTGGTGGGCAATTTGTGCCAGAGGCACTGGGCTGCTCTCCCTGAG 49117
Qy 6000 GATCCTCTGCCTCACTACCTCCATCCAGGGCACTTGCCTCAGCCAAAGACCTACTATCCGGATGA 6059
Db 49118 GATCCTCTGCCTCACTACCTCCATCCAGGGCACTTGCCTCAGCCAAAGACCTACTATCCGGATGA 49177
Qy 6060 ACACATCTACCCCTTACATGATACCTTGGTGGCTACCACTGTCCGCAACCGCAATGTGCGGGA 6119
Db 49178 ACACATCTACCCCTTACATGATACCTTGGTGGCTACCACTGTCCGCAACCGCAATGTGCGGGA 49237
Qy 6120 AGCCCTCGAGGCTGGGGGACAGGCCACTGTCTATCCAGAGTGGAGTCCCTACTA 6179
Db 49238 AGCCCTCGAGGCTGGGGGACAGGCCACTGTCTATCCAGAGTGGAGTCCCTACTA 49297
Qy 6180 GGGCTCGAGCTCTCTCTTCTTCCCTTCCATCACTAGTTTCCAACCACTCTCTCCAGGACT 6239
Db 49298 GGGCTCGAGCTCTCTCTTCTTCCCTTCCATCACTAGTTTCCAACCACTCTCTCCAGGACT 49357
Qy 6240 GAGGCTGGCTTCCAGGCCCATCCCTTCCATCCAGTCCCTAGGCAAGCAAGGCCAGCA 6299
Db 49358 GAGGCTGGCTTCCAGGCCCATCCCTTCCATCCAGTCCCTAGGCAAGCAAGGCCAGCA 49417
Qy 6300 TTACCCAGGAGGTAGGACCTGATTAAGGTGCACATCTTTCCTCCCTCCCTCTCTCTCT 6359
Db 49418 TTACCCAGGAGGTAGGACCTGATTAAGGTGTACATCTTTCCTCTCTCTCTCTCTCT 49477
Qy 6360 CCTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAATCTTAAACCACCATCATCCAG 6419
Db 49478 CCTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAATCTTAAACCACCATCATCCAG 49537
Qy 6420 CAGTGGGACTTCCACCTCGGCCCATGCCCCCTCTCTCATCTTCTGCTTCTCTCTGG 6479
Db 49538 CAGTGGGACTTCCACCTCGGCCCATGCCCCCTCTCTCATCTTCTGCTTCTCTCTGG 49597
Qy 6480 GCTCAGCCAGACAGCATCTTTTGCAGTGAGAGCCCACTACTCCCCAGCCCTCTGGG 6539
Db 49598 GCTCAGCCAGACAGCATCTTTTGCAGTGAGAGCCCACTACTCCCCAGCCCTCTGGG 49657
Qy 6540 GCTCCATCCCCCGCAGGCTCCTGGGGTACCCCGAT - GGTGAGACCCCTTTCAGACCT 6598
Db 49658 GCTCCATCCCCCGCAGGCTCCTGGGGTACCCCGATGGGTGAGACCCCTTTCAGACCT 49717
Qy 6599 ACAGAGCCCACTGCTCAGACTACAATCTGCGGGAAGCAAGAGAGAGATCTACAA 6658
Db 49718 ACAGAGCCCACTGCTCAGACTACAATCTGCGGGAAGCAAGAGAGAGATCTACAA 49777
Qy 6659 GGAGTCTCTTGAAGTAGCCAAATGATGTATCCCAACCTGCTGAAGAGGAGGAGGAGCTT 6718
Db 49778 GGAGTCTCTTGAAGTAGCCAAATGATGTATCCCAACCTGCTGAAGAGGAGGAGGAGCTT 49837
Qy 6719 GCTGGAGGGGGGAGGAGGCGGGGAGCAAGCCAGGTGAAGGCTGGAUCCCTCAG 6778
Db 49838 GCTGGAGGGGGGAGGAGGCGGGGAGCAAGCCAGGTGAAGGCTGGAUCCCTCAG 49897
Qy 6779 CTTGTGTCACCCCTCCACCTGGAGAGGGCTTCCCTCCACAGGCCATGGGGCTGCAATG 6838
Db 49898 CTTGTGTCACCCCTCCACCTGGAGAGGGCTTCCCTCCACAGGCCATGGGGCTGCAATG 49956
Qy 6839 TACGGGATTAGGGATGGCAGGAGGAGGAGGTGGCC - CTGAGCAGACAGCTATGTTCCTTTT 6897
Db 49957 TACGGGATTAGGGATGGCAGGAGGAGGAGGTGGCCACTGAGCAGACAGCTATGTTCCTTTT 50016
Qy 6898 GCTATACTGAGGTCTCTGGGCCACAGTTTGGAGC 6930
Db 50017 GCTATACTGAGGTCTCTGGGCCACAGTTTGGAGC 50049

AP000575

LOCUS 73431 bp DNA HTG 30-MAY-2000
 DEFINITION Homo sapiens chromosome 11 clone XXpl-95F6 map 11ql13, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.
 ACCESSION AP000575
 VERSION AP000575.3 GI:8118783
 KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
 SOURCE Homo sapiens DNA, clone:XXpl-95F6.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 73,431 genomic DNA of 11ql13

Published Only in Database (1999) In press

REFERENCE

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (08-OCT-1999) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,

Japan (E-mail:hattori@psc.riken.go.jp,

URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,

Fax:81-42-778-9924)

On May 31, 2000 this sequence version replaced gi:6997467.

COMMENT

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@psc.riken.go.jp

----- Project Information

Center project name: HumDraft11

Center clone name: XXpl-95F6

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 66753 bases at least Q40

Consensus quality: 69862 bases at least Q30

Consensus quality: 71369 bases at least Q20

Insert size: 72031; sum-of-contigs

Quality coverage: 4.73x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 15 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

1 10174 contig of 10174 bp in length
 10275 20667 contig of 10393 bp in length
 20768 29453 contig of 8686 bp in length
 29554 37145 contig of 7592 bp in length
 37246 41949 contig of 4704 bp in length
 42050 47344 contig of 5295 bp in length
 47445 52652 contig of 5208 bp in length
 52753 56084 contig of 3332 bp in length
 56185 60854 contig of 4670 bp in length
 60955 63499 contig of 2545 bp in length
 63600 66501 contig of 2902 bp in length
 66602 68545 contig of 1944 bp in length
 68646 70748 contig of 2103 bp in length
 70849 72288 contig of 1440 bp in length
 72389 73431 contig of 1043 bp in length

Sequence updated (12-Oct-1999)

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 10174: contig of 10174 bp in length
 10175 10274: gap of 100 bp
 10275 20667: contig of 10393 bp in length
 20668 20767: gap of 100 bp
 20768 29453: contig of 8686 bp in length
 29454 29553: gap of 100 bp
 29554 37145: contig of 7592 bp in length
 37146 37245: gap of 100 bp
 37246 41949: contig of 4704 bp in length
 41950 42049: gap of 100 bp
 42050 47344: contig of 5295 bp in length
 47345 47444: gap of 100 bp
 47445 52652: contig of 5208 bp in length
 52653 52752: gap of 100 bp
 52753 56084: contig of 3332 bp in length
 56085 56184: gap of 100 bp
 56185 60854: contig of 4670 bp in length
 60855 60954: gap of 100 bp
 60955 63499: contig of 2545 bp in length
 63500 63599: gap of 100 bp
 63600 66501: contig of 2902 bp in length
 66502 66601: gap of 100 bp
 66602 68545: contig of 1944 bp in length
 68546 68645: gap of 100 bp
 68646 70748: contig of 2103 bp in length
 70749 70848: gap of 100 bp
 70849 72288: contig of 1440 bp in length
 72289 72388: gap of 100 bp
 72389 73431: contig of 1043 bp in length.

FEATURES

Source

1. 73431
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11ql13"
 /clone="XXpl-95F6"

misc_feature 1..10174
 /note="assembly_fragment"
 misc_feature 10275..20667
 /note="assembly_fragment"
 misc_feature 20768..29453
 /note="assembly_fragment"
 misc_feature 29554..37145
 /note="assembly_fragment"
 misc_feature 37246..41949
 /note="assembly_fragment"
 misc_feature 42050..47344
 /note="assembly_fragment"
 misc_feature 47445..52652
 /note="assembly_fragment"
 misc_feature 52753..56084
 /note="assembly_fragment"
 misc_feature 56185..60854
 /note="assembly_fragment"
 misc_feature 60955..63499
 /note="assembly_fragment"
 misc_feature 63600..66501
 /note="assembly_fragment"
 misc_feature 66602..68545
 /note="assembly_fragment"
 misc_feature 68646..70748
 /note="assembly_fragment"
 misc_feature 70849..72288
 /note="assembly_fragment"
 misc_feature 72389..73431
 /note="assembly_fragment"

BASE COUNT 17276 a 18750 c 19264 g 16741 t 1400 others
 ORIGIN

[illegible]

QY	2064	TTGTGGGGACAATAAGGCTCTGCAGTCTCGGCTGAGGGTGCTCACCAACAAAGAGGGG	2123
Db	69666	TTGTGGGGACAATAAGGCTCTGCAGTCTCGGCTGAGGGTGCTCACCAACAAAGAGGGG	69725
QY	2124	AAGCCGGTGACGAGAGGCTGAAGAGGTTGGGAACGAGGGAGCTGTCGCTGTGTGTCGGGG	2183
Db	69726	AAGCCGGTGACGAGAGGCTGAAGAGGTTGGGAACGAGGGAGCTGTCGCTGTGTGTCGGGG	69785
QY	2184	CGGTTGGAACCTTAGCGGACCCTGGGAGAGGCTGCCCGCGCCGAACCTGCCCGACCCCTCC	2243
Db	69786	CGGTTGGAACCTTAGCGGACCCTGGGAGAGGCTGCCCGCGCCGAACCTGCCCGACCCCTCC	69845
QY	2244	CTCCCCCGCTTGCCCTTCAGAGCGCCGCCCANCGCCGCCCATGGGCTGAAGGCGG	2303
Db	69846	CTCCCCCGCTTGCCCTTCAGAGCGCCGCCCANCGCCGCCCATGGGCTGAAGGCGG	69905
QY	2304	CCAGAAGACGCTGTTCGCCGTGCGCTCCAATCGAGAGCTGGTGGCGCCTGTTGCTGGCG	2363
Db	69906	CCAGAAGACGCTGTTCGCCGTGCGCTCCAATCGAGAGCTGGTGGCGCCTGTTGCTGGCG	69965
QY	2364	AGCTGGGCGGAGAGAGCGGACCTGGTGTCTCTTTTCCCTGGTGTGGCTTCGTGGAGC	2423
Db	69966	AGCTGGGCGGAGAGAGCGGACCTGGTGTCTCTTTTCCCTGGTGTGGCTTCGTGGAGC	70025
QY	2424	ATTTTCTGGCTGTCAAACCGGCTCATCCCTACCAAGCTTCCCAGAGCTCACTTCCAAGCCCA	2483
Db	70026	ATTTTCTGGCTGTCAAACCGGCTCATCCCTACCAAGCTTCCCAGAGCTCACTTCCAAGCCCA	70085
QY	2484	GCCCGCCCGACCCGCGCTGGCGGCTCACTACTTTCCGCTGGCGGACCTGCTCATCA	2543
Db	70086	GCCCGCCCGACCCGCGCTGGCGGCTCACTACTTTCCGCTGGCGGACCTGCTCATCA	70145
QY	2544	TGCGCGCCCTCTATGCGCGCTTTCACGCGCCAGATCCGAGGCGCGCTGCACCTGTCCCTCT	2603
Db	70146	TGCGCGCCCTCTATGCGCGCTTTCACGCGCCAGATCCGAGGCGCGCTGCACCTGTCCCTCT	70205
QY	2604	ATCTCTGAGAAGGGGTGCTCTCCAGCCGTGAGCTGGTGAAGAAGCTCTCCGATGTCAAT	2663
Db	70206	ATCTCTGAGAAGGGGTGCTCTCCAGCCGTGAGCTGGTGAAGAAGCTCTCCGATGTCAAT	70265
QY	2664	GGAACAGCTCAGCGCGCTCTACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAGCT	2723
Db	70266	GGAACAGCTCAGCGCGCTCTACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAGCT	70325
QY	2724	TCATCACAGTTTGGAGCCCGAGTGTGGGAATCTTTATCCATGACCCACTTCTTCAAAAACC	2783
Db	70326	TCATCACAGTTTGGAGCCCGAGTGTGGGAATCTTTATCCATGACCCACTTCTTCAAAAACC	70385
QY	2784	CTCCATGCTTTACAGAACCTTTTAAGAACTGTAAAGCTTGTGAGGTTTCGGCAGCTGTGA	2843
Db	70386	CTCCATGCTTTACAGAACCTTTTAAGAACTGTAAAGCTTGTGAGGTTTCGGCAGCTGTGA	70445
QY	2844	TTTTTCTCTTTTCAGCTTGGAAACTGAAGCCACAGAGGGGAAATGATATGCCAAGTCA	2903
Db	70446	TTTTTCTCTTTTCAGCTTGGAAACTGAAGCCACAGAGGGGAAATGATATGCCAAGTCA	70505
QY	2904	CACAGGCATGCGAGGGCTGGAAGTGAAGCCTGATCACTTGCTGCCAATATCAACCTC	2963
Db	70506	CACAGGCATGCGAGGGCTGGAAGTGAAGCCTGATCACTTGCTGCCAATATCAACCTC	70565
QY	2964	ACCTCTGCCCTCTAGCAGCCGCCACCTTGCCACTGAACAGCTACAGAGTTCTAAGCAT	3023
Db	70566	ACCTCTGCCCTCTAGCAGCCGCCACCTTGCCACTGAACAGCTACAGAGTTCTAAGCAT	70625
QY	3024	GAGACACAGAGGGCGGACAGATTTAGGGGCAAGAAGATCAAAATTTGGGCTGCATTTGA	3083
Db	70626	GAGACACAGAGGGCGGACAGATTTAGGGGCAAGAAGATCAAAATTTGGGCTGCATTTGA	70685
QY	3084	GCGAGTTAAACAAAAATAATGGCTATCAAGATTTTTTTTTTTTTTTTTTTTTTTTTCAGACAGG	3143
Db	70686	GCGAGTTAAACAAAAATAATGGCTATCAAGATTTTTTTTTTTTTTTTTTTTTTTTTCAGACAGG	70745
QY	3144	TCCTACTCTGTCCCCCAGGCTGGAGTGGCTGTGATCATGGCTCACTGGCAGGCTCAG	3203

108640	115961	contig of	7322 bp in length
116062	120258	contig of	4197 bp in length
120359	125359	contig of	5001 bp in length
125460	130186	contig of	4727 bp in length
130287	134000	contig of	4114 bp in length
134501	139064	contig of	4564 bp in length
139165	143906	contig of	4742 bp in length
144007	148280	contig of	4274 bp in length
148381	151006	contig of	2626 bp in length
151107	153748	contig of	2642 bp in length
153849	155555	contig of	1707 bp in length
155656	157198	contig of	1543 bp in length
157299	159872	contig of	2574 bp in length
159973	161499	contig of	1527 bp in length
161600	162703	contig of	1104 bp in length
162804	163847	contig of	1044 bp in length
163948	165121	contig of	1174 bp in length
165222	166269	contig of	1048 bp in length
* NOTE: This is a 'working draft' sequence. It currently consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			
1	28395:	contig of	28395 bp in length
28396	28495:	gap of	100 bp
28496	43446:	contig of	14951 bp in length
43447	43546:	gap of	100 bp
43547	58083:	contig of	14537 bp in length
58084	58183:	gap of	100 bp
58184	66868:	contig of	8685 bp in length
66869	66968:	gap of	100 bp
66969	78574:	contig of	9606 bp in length
78575	76674:	gap of	100 bp
76675	86418:	contig of	9744 bp in length
86419	86518:	gap of	100 bp
86519	94044:	contig of	7526 bp in length
94045	94144:	gap of	100 bp
94145	100790:	contig of	6646 bp in length
100791	100890:	gap of	100 bp
100891	108539:	contig of	7649 bp in length
108540	108639:	gap of	100 bp
108640	115961:	contig of	7322 bp in length
115962	116061:	gap of	100 bp
116062	120258:	contig of	4197 bp in length
120259	120358:	gap of	100 bp
120359	125359:	contig of	5001 bp in length
125360	125459:	gap of	100 bp
125460	130186:	contig of	4727 bp in length
130187	130286:	gap of	100 bp
130287	134400:	contig of	4114 bp in length
134401	134500:	gap of	100 bp
134501	139064:	contig of	4564 bp in length
139065	139164:	gap of	100 bp
139165	143906:	contig of	4742 bp in length
143907	144006:	gap of	100 bp
144007	148280:	contig of	4274 bp in length
148281	148380:	gap of	100 bp
148381	151006:	contig of	2626 bp in length
151007	151106:	gap of	100 bp
151107	153748:	contig of	2642 bp in length
153749	153848:	gap of	100 bp
153849	155555:	contig of	1707 bp in length
155556	155655:	gap of	100 bp
155656	157198:	contig of	1543 bp in length
157199	157298:	gap of	100 bp
157299	159872:	contig of	2574 bp in length
159873	159972:	gap of	100 bp
159973	161499:	contig of	1527 bp in length
161500	161599:	gap of	100 bp
161600	162703:	contig of	1104 bp in length

Query Match 21.6%; Score 1982.2; DB 82; Length 166269;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 2054; Conservative 0; Mismatches 128; Indels 5; Gaps 5;

[illegible]

Db 148356 NNNNNNNNNNNNNNNNNNNNNNNC 148382

```
RESULT 11
HSU93236
LOCUS          HSU93236       2772 bp      mRNA          PRI          19-APR-1997
DEFINITION     Human menin (MEN1) mRNA, complete cds.
ACCESSION      U93236
VERSION        U93236.1  GI:1945386
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               1 (bases 1 to 2772)
               Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.-E.,
               Collins,F.S., Emmert-Buck,M.R., Debelenko,L.V., Zhuang,Z.,
               Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A.,
               Weisenmann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S.,
               Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.,
               Positional cloning of the gene for multiple endocrine
               neoplasia-type 1
               Science 276 (5311), 404-407 (1997)
JOURNAL
MEDLINE        97258940
REFERENCE      2 (bases 1 to 2772)
AUTHORS        Collins,F.S.
TITLE          Direct Submission
JOURNAL        Submitted (13-MAR-1997) National Human Genome Research Institute,
               Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD
               20892, USA
FEATURES
source         Location/Qualifiers
               1..2772
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="11"
               /map="11q13"
               1..87
               /gene="MEN1"
               /number=1
               1..2772
               /gene="MEN1"
               /number=1
               88..555
               /gene="MEN1"
               /number=2
               111..1943
               /gene="MEN1"
               /codon_start=1
               /product="menin"
               /protein_id="AAC51228.1"
               /db_xref="GI:1945387"
               /translation="MGLKAAQKTLPLRSIDDVRLFAAEELGPEEDLVLLSLVLGFG
               EHLFVNRVPTNPVETLFGPSPPADPPGGTLTFPVADLSI IAAFYARF"AOIRGAVFA
               LSLYPREGGVSSRELKPKSDVILNLSRSYFKDRAHIQSLFSFIQTCLDSSGSAFA
               VVGAALGLRDVHLALSEBHAVVYVPGNGEQTAEVTWGKGNEDRGQTVNAGVAER
               SWILKGSIMRCMKMEVAMVCANPISDIHDSLELIQLOQRKLWLLYDIGHRLRY
               PMALQWADATLEETLPGRPDPTLTLYHKGIASAKTYIRDEHIYPIYMLAGYHCRNRY
               REALQWADATYIQDYNCREDEEYKEEFVANDVIPNLKKAASLLAEGERGE
               QSQGTQSGSALADPECFALHLRFYDGIKWEBSGPTPLHVGWATFLVQSLGFEFG
               VRQKRVTSREAAAEAEPEWBEAREGRRRRGPRRESKPEPPPKPKKPADLKGLTGTG
               GAVSGPRKPGTVAGTARGPEGSTAQVAPASPPPEGPVLTFOSEMKMGKNKELL
               ATKINSRAIKLQLTAQSQVQMKKQVSTFSDYTLSPFLKQRKGL"
               175
               /gene="MEN1"
               /standard_name="L22R"
               /note="Lys22Arg"
               /replace="g"
               357..360
               /gene="MEN1"
               /standard_name="357del4"
               /note="4 bp deletion in kindred 0176"
               /replace=""
               416
               /gene="MEN1"
               /standard_name="MEN1"
               /note="MEN1"
```

/standard_name="416delC"
/note="1 bp deletion"
/replace=""
464..466
/gene="MEN1"
/standard_name="K119del"
/note="3 bp deletion"
/replace=""
512
/gene="MEN1"
/standard_name="512delC"
/note="1 bp deletion"
/replace=""
556..764
/gene="MEN1"
/number=3
704
/gene="MEN1"
/standard_name="W198X"
/note="Trp198stop"
/replace="a"
735..738
/gene="MEN1"
/standard_name="735del4"
/note="4 bp deletion"
/replace=""
765..893
/gene="MEN1"
/number=4
894..934
/gene="MEN1"
/number=5
935..1022
/gene="MEN1"
/number=6
1023..1159
/gene="MEN1"
/number=7
1132
/gene="MEN1"
/standard_name="1132delG"
/note="1 bp deletion"
/replace=""
1160..1295
/gene="MEN1"
/number=8
1197..1199
/gene="MEN1"
/standard_name="E363del"
/note="3 bp deletion"
/replace=""
1296..1460
/gene="MEN1"
/number=9
1416
/gene="MEN1"
/standard_name="W436R"
/note="Trp436Arg"
/replace="c"
1417
/gene="MEN1"
/standard_name="W436X"
/note="Trp436stop"
/replace="a"
1461..2772
/gene="MEN1"
/number=10
1689
/gene="MEN1"
/standard_name="R527X"
/note="Arg527stop"
/replace="t"
577 a 909 c 779 g 507 t

ORIGIN

Query Match 14.3%; Score 1312.2; DB 97; Length 2772;
Best Local Similarity 96.5%; Pred. No. 3.2e-268; Mismatches 48; Indels 0; Gaps 0;
Matches 1341; Conservative 0;
Qy 7499 GGATGGCCAGAGCAGGTCCTGGAGTTCCAGCCACTGCGCGGCAACCTTGGCTCTCACCTT 7558
Db 1382 GGAGGCGCAGTCCCACGCTGTGTGCACGTGGGTGGGCCACCTTTCTTTGTGCAGTCCCT 1441
Qy 7559 GCTCTCCCCACTGGCCCCAGGTGGCGGAGAGGTGGCATAGTGAGCCGAGAGCGGAGGC 7618
Db 1442 AGCCCTTTTGAAGGACAGGTGGCGCAGAGGTGGCATAGTGAGCCGAGAGCGGAGGC 1501
Qy 7619 GGCGAGGCGCAGGAGCGCTGGGGCGGAGGAGCCCGGGAAGCGCGGCGGCGGCGGCCACG 7678
Db 1502 GGCGAGGCGCAGGAGCGCTGGGGCGGAGGAGCCCGGGAAGCGCGGCGGCGGCGGCCACG 1561
Qy 7679 GCGGAGTCCCAAGCCAGAGGAGCGCGCGCGCCCAAGAACCCAGCACCTGTG;ACAAGGGCCT 7738
Db 1562 GCGGAGTCCCAAGCCAGAGGAGCGCGCGCGCCCAAGAACCCAGCACCTGTGACAAGGGCCT 1621
Qy 7739 GGGCACGGCCAGGTCAGTGTTCAGGACCCCGCGGGAAGCCTCTGGGACTGTGCTGG 7798
Db 1622 GGGCACGGCCAGGTCAGTGTTCAGGACCCCGCGGGAAGCCTCTGGGACTGTGCTGG 1681
Qy 7799 CACAGCCGAGGCGCTGAAGGTGGCAGCACGGCTCAGTGCCAGCACCCCGCAGCATCACC 7858
Db 1682 CACAGCCGAGGCGCTGAAGGTGGCAGCACGGCTCAGTGCCAGCACCCCGCAGCATCACC 1741
Qy 7859 ACCCGCGGAGGTCAGTGTCTACTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCT 7918
Db 1742 ACCCGCGGAGGTCAGTGTCTACTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCT 1801
Qy 7919 GCTGTTGGCCACCAAGATCAACTCGAGGCGCATCAAGTGCACACTCAGGCGACAGTCGCA 7978
Db 1802 GCTGTTGGCCACCAAGATCAACTCGAGGCGCATCAAGTGCACACTCAGGCGACAGTCGCA 1861
Qy 7979 AGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTTCTCTCAA 8038
Db 1862 AGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTTCTCTCAA 1921
Qy 8039 GCGCAGCGCAAAAGCGCTCTGAACCTACTGGGAGCTTCGGACCGCTTTGGGAGCCACAGC 8098
Db 1922 GCGCAGCGCAAAAGCGCTCTGAACCTACTGGGAGCTTCGGACCGCTTTGGGAGCCACAGC 1981
Qy 8099 TCCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCC 8158
Db 1982 TCCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCC 2041
Qy 8159 TCACCTCTACCCAAACCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACCAT 8218
Db 2042 TCACCTCTACCCAAACCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACCAT 2101
Qy 8219 TTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCGGTTAGAAATATAGTCTCTTTCTGCC 8278
Db 2102 TTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCGGTTAGAAATATAGTCTCTTTCTGCC 2161
Qy 8279 AGAATCCAGCGCGCCCAATGAAACCTCACGCTGGTCTTAATTTACCAAGTCTTTAAAGGC 8338
Db 2162 AGAATCCAGCGCGCCCAATGAAACCTCACGCTGGTCTTAATTTACCAAGTCTTTAAAGGC 2221
Qy 8339 CCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAACCTTT 8398
Db 2222 CCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAACCTTT 2281
Qy 8399 ACTCAGGGCTCTCCCGAGCTTGTAGGAGCTGAGGTTTTCACCCCTTAACCCAGGAGCACA 8458
Db 2282 ACTCAGGGCTCTCCCGAGCTTGTAGGAGCTGAGGTTTTCACCCCTTAACCCAGGAGCACA 2341
Qy 8459 GGTCCCACTCCAGCGCGGAGGCTTAGGAGCTCAGCCCTTAGGAGTATATTTCCGCA 8518

Db 960 CTCACCTCTACCAACACCTAGTTCCCGTCCCGAGTACAGTCTGTATCAAAACCCACGA 1019
Qy 8218 TTTTCTCCAGTCAAGACCCAGGGCTCTGCCAGTCTGTAGATATAGGCTCTCTTC 8277
Db 1020 TTTTCTCCAGTCAAGACCCAGGGCTCTGCCAGTCTGTAGATATAGGCTCTCTTC 1079
Qy 8278 CAGATCCACGCGGCCAATGGAACCTCAGCTGGTCTCTTAATACCAAGCTTTTAAAG 8337
Db 1080 CAGATCCACGCGGCCAATGGAACCTCAGCTGGTCTCTTAATACCAAGCTTTTAAAG 1139
Qy 8338 CCCAGCCCTAGAACCCCAAGCTCTCTCGAAGCCGCTACCTAGACCCAGCAAGCT 8397
Db 1140 CCCAGCCCTAGAACCCCAAGCTCTCTCGAAGCCGCTACCTAGACCCAGCAAGCT 1199
Qy 8398 TACTCAGGCTCTCCACAGCTTGTAGGAGCTGAGTTCACCCCTTAACCCAGGAGCAGC 8457
Db 1200 TACTCAGGCTCTCCACAGCTTGTAGGAGCTGAGTTCACCCCTTAACCCAA -GGAGCAGC 1258
Qy 8458 AGGTCCACCTCCAGCCCGGGAGCCTAGGACCACTAGCCCTCAGCCCTAGGAGTATATTCGCG 8517
Db 1259 AGGTCCACCTCCAGCCCGGGAGCCTAGGACCACTAGCCCTCAGCCCTAGGAGTATATTCGCG 1317
Qy 8518 ACTTCAGATTCATATCTTGGGAATCAAGCTCCCTGCCCAATAACTTCAGTCTCTGC 8577
Db 1318 ACTTCAGATTCATATCTTGGGAATCAAGCTCCCTGCCCAATAACTTCAGTCTCTGC 1377
Qy 8578 TTCCAGATTTGGAATCCTAGTTTCTCTCTCTCGATCCGAGCTCGGACACAAAC 8637
Db 1378 -TCCAGATTTGGAATCCTAGTTTCTCTCTCTCGATCCGAGCTCGGACACAAAC 1436
Qy 8638 TCCGCCCCACGCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCGCG 8697
Db 1437 TCCGCCCCACGCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCGCG 1496
Qy 8698 ATCAGAGCAGGACCTCCCTCTCGACCTCTGGACCTCCAGAGGTCAGCCGATCTCG 8757
Db 1497 ATCAGAGCAGGACCTCCCTCTCGACCTCTGGACCTCCAGAGGTCAGCCGATCTCG 1556
Qy 8758 GAGCATCCCGAGGAAATCTCGACAGGGTTAGGAGTGGGTGACAGGCTGTATCTCT 8817
Db 1557 GAGCATCCCGAGGAAATCTCGAGA -GGTTAGGAGTGGGTGACAGGCTGTATCTCT 1615
Qy 8818 CTTGTTTGTACATAGATTTATTTTTCAGTTCAGTTCAGGAAAGATGATATTTGTTAAAA 8877
Db 1616 CTTGTTTGTACATAGATTTATTTTTCAGTTCAGTTCAGGAAAGATGATATTTGTTAAAA 1675
Qy 8878 AATATATAA 8887
Db 1676 AAAAAAAAA 1685

RESULT 13
BC002544
LOCUS BC002544 2633 bp mRNA PRI 16-MAR-2001
DEFINITION Homo sapiens, Similar to multiple endocrine neoplasia I, clone
MGC:1554, mRNA, complete cds.
ACCESSION BC002544
VERSION BC002544.1 GI:12803440
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2633)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
CONTACT: Robert Strausberg, Ph.D.
COMMENT Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Shvchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Detrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 4 Row: f Column: 7.

FEATURES

Source 1..2633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:1554 IMAGE:3138471"
/tissue_type="Placenta, Chorioncarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
78..1805
CDS
/codon_start=1
/product="Similar to multiple endocrine neoplasia I"
/protein_id="AAH02544.1"
/db_xref="GI:12803441"
/translation="MGLKAAQKTLPLRSIDVVVLEAELGREEPDLVLLSILVGFV
EHLAVNRVPTNNVPELTFQSPADPPGLTYEPVADLSIARLYARFNAQJRGAVD
LSLYPRGGVSSRELKVKSDVWNSLSRYSFKDRAHISLFSFTGTGLDSSGVAFA
VVGACQALGRDLVHLSLSEDAHWSMLYKGSYMRCKRMEVAFVCAINPSIDLITDS
LELILOQLKLLWLYDLGHLERYPMALGNLADLELEPTGPDPLTYIKIGIASAKT
VYRDEHYPMYLAGYHCRNRNVRQESQGSALQDPECFAHLLRYVDCIKWEEG
DVIPNLKAEASLLLEAGEERPEQSGSALQDPECFAHLLRYVDCIKWEEG
PTPLVHVWATFLVQSLRGEGVQRVRYVREAAEAEEPEHGEARERRGRPR
ESKPEPPPPKPAIDKGLGTGGVSGPPRPKPGTAVGTARGPEGEAQVAPPAAS
PPPEPVLTFTQSEKMKMKELLATKINSSAIKQLTAQSQVOMKKOKVSTPSDTLS
FLKQRKGL"

BASE COUNT 548 a 865 c 729 g 491 t
ORIGIN

Query Match 13.6%; Score 1251.6; DB 91; Length 2633;
Best Local Similarity 96.1%; Pred. No. 2.3e-255;
Matches 1336; Conservative 0; Mismatches 49; Indels 5; Gaps 5;
Qy 7499 GGATGGCCAGACAGGCTCTGGAGTTCAGCCACTGGCCGCAACCTTGTCTCACCTT 17558
Db 1244 GGAGGGCAGTCCAGCCGCTGTGTCAGCTGGGCTGGGCCACCTTTCTTGTGACGTCCT 1303
Qy 7559 GCTCTCTCCACTGGCCGAGGTCGCGCAGAGGTGCGCATAGTGAGCCGAGAGGCTGAGCG 7618
Db 1304 AGGCCGTTTGGAGGACAGGTGCGGCAGAGGTGCGCATAGTGAGCCGAGAGGCTGAGCG 1363
Qy 7619 GGCCGAGCCGAGGAGCGTGGGGCGAGAGCCCGGGAAGCCGCGGCGGGGGCCCCAGG 7678
Db 1364 GGCCGAGCCGAGGAGCGTGGGGCGAGAGCCCGGGAAGCCCGGCGGGGGCCCCAGG 1423
Qy 7679 GCGGAGTCCAGCAGAGGAGCCCGCCCGCCCAAGAGCCAGCAGCTGGACAAGGGCT 7738
Db 1424 GCGGAGTCCAGCAGAGGAGCCCGCCCGCCCAAGAGCCAGCAGCTGGACAAGGGCT 1483
Qy 7739 GGGCACCGCCAGGCTGTCAGTGTACAGAGCCCGCCCGGAGCCCTCTGGGACTGTCCCTGG 7798
Db 1484 GGGCACCGCCAGGCTGTCAGTGTACAGAGCCCGCCCGGAGCCCTCTGGGACTGTCCCTGG 1543

QY 7799 CACAGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCAGCCGACATCAC 7858
Db 1544 CACAGCCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCAGCCGACATCAC 1603
-QY 7859 ACCGCGGAGGCTCAGTCTCTCTTCCAGAGTGAAGATGAAGGATGAAGGAGCT 7918
Db 1604 ACCGCGGAGGCTCAGTCTCTCTTCCAGAGTGAAGATGAAGGATGAAGGAGCT 1663
QY 7919 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGACAGTCGCA 7978
Db 1664 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGACAGTCGCA 1723
QY 7979 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCAA 8038
Db 1724 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCAA 1783
QY 8039 GCGGACGCGCAAGGCTCTGAACCTACTGGGACCTTCGGACCGCTTGTGGGACCCAGCG 8098
Db 1784 GCGGACGCGCAAGGCTCTGAACCTACTGGGACCTTCGGACCGCTTGTGGGACCCAGCG 1843
QY 8099 TCCG-CCTTAGTCCGCCAACTCTGAGCCATGTCTGCCGCCAGCCCAAGGGGACAGCG 8157
Db 1844 TCCGCCCTTAGTCCGCCAACTCTGAGCCATGTCTGCCGCCAGCCCAAGGGGACAGCG 1903
QY 8158 CTCACCTCTACCCAAACCTAGGTTCGCCGTCCCGGTACAGTCTGTATCAAAACCCACGA 8217
Db 1904 CTCACCTCTACCCAAACCTAGGTTCGCCGTCCCGGTACAGTCTGTATCAAAACCCACGA 1963
QY 8218 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGCTGTAGATATAGTCTCTCTCC 8277
Db 1964 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGCTGTAGATATAGTCTCTCTCC 2023
QY 8278 CAGATCCCGAGCGGCAATGGAACCTCAGCGTGGTCTTAATACCACTCTTTAAAGG 8337
Db 2024 CAGATCCCGAGCGGCAATGGAACCTCAGCGTGGTCTTAATACCACTCTTTAAAGG 2083
QY 8338 CCAGCCCTTAGAACCCAAAGCTCTCTCGGAACCGCTCACCTAGAGCAGACCAACGT 8397
Db 2084 CCAGCCCTTAGAACCCAAAGCTCTCTCGGAACCGCTCACCTAGAGCAGACCAACGT 2143
QY 8398 TACTAGGCTCTCCAGCTGTAGAGCTGAGTTTACCTTACCTTAAACCAAGGAGCAC 8457
Db 2144 TACTAGGCTCTCCAGCTGTAGAGCTGAGTTTACCTTAAACCAAGGAGCAC 2202
QY 8458 AGTCCACCTCCAGCCGCGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCCGC 8517
Db 2203 AGTCCACCTCCAGCCGCGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCCGC 2261
QY 8518 ACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTGCCCAAAATTAATTCAGTCTGC 8577
Db 2262 ACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTGCCCAAAATTAATTCAGTCTGC 2321
QY 8578 TTCCAGAAATTTGAAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8637
Db 2322 -TCCAGAAATTTGAAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2380
QY 8638 TCCGCCCCAGCCTATGACATCTCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8697
Db 2381 TCCGCCCCAGCCTATGACATCTCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2440
QY 8698 ATCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCG 8757
Db 2441 ATCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCG 2500
QY 8758 GAGCATCCCGGAGGAATCTGCAGAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTT 8817
Db 2501 GAGCATCCCGGAGGAATCTGCAGAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTT 2559
QY 8818 CCGTTTTGTACATATATTTTTCAGTTCCCAAGAAATGATCAATATTTTGTATAA 8877
Db 2560 CCGTTTTGTACATATATTTTTCAGTTCCCAAGAAATGATCAATATTTTGTATAA 2619
QY 8878 AAAATATAA 8887

Db 2620 AAAAAAAAAA 2629
|||||
RESULT 14
HSURF
LOCUS
DEFINITION
H.sapiens mRNA 3'-region (unknown function). PRI 03-JUL-1998
ACCESSION Y12338
VERSION Y12338.1 GI:2217971
KEYWORDS
SOURCE human.
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES
Location/Qualifiers
1. .1837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11q13"
<1..1027
/function="unknown"
/note="unnamed protein product"
/codon_start=2
/protein_id="CAA73007.1"
/db_xref="GI:2217972"
/translation="LGHLEPYDMALGNLADLELEPTGCRDPLTLYHKGIASAKTY
RDEHLYPYMYLAGYHCRNRNVREALQAWADATVTDYNYCREDEIYKEFEVANDV
IPNLKAEASLLEAGEERPEQSGSALQDQPECFALLRLFYDGLCKWEESGPT
PVLHKGWATFLVLSIGREFGQVQKRVISREAAEAEEPEWGEAREGRRGPRRES
KPEEPPPKPALDKGLGTGAGVSPRPKPPGTAGTARGGEGTAGVPAPASPP
PGPVLTTFQSEKMKMKELLVATKINSSAIKQLTAQSOVOMKKOKVSTPSDTLSFL
KQRKGL"
BASE COUNT 412 a 599 c 500 g 326 t
ORIGIN
Query Match 13.5%; Score 1242.4; DB 97; Length 1837;
Best Local Similarity 96.3%; Pred. No. 2.1e-253;
Matches 1325; Conservative 0; Mismatches 46; Indels 5; Gaps 5;
QY 7499 GGATGCCAGAGCAGGCTCCTGGAGTTCAGCCACTGGCGGCAACCTTGCTCTCACTT 7558
Db 466 GGAGGCGAGTCCACGCCCTGTCTGTGCACGTGGGCTGGCCACCTTCTTGTGAGTCCCT 525
QY 7559 GCTCTCCCACTGGCCCGGAGAGGTTGCGCATAGTCCGAGCGGAGCGGAGG 7618
Db 526 AGCCCGCTTTGAGGGGACAGGTGCGGAGAGGTGCGCATAGTCCGAGCGGAGCGGAGG 585
QY 7619 GGCGGAGGCGGAGGAGCGCTGGGGCGGAGGAGCGCGGGGCGGCGGCGGCGG 7678
Db 586 GGCGGAGGCGGAGGAGCGCTGGGGCGGAGGAGCGCGGGGCGGCGGCGGCGGCGG 645
QY 7679 GCGGAGTCCAAAGCCAGAGAGCGCCCGCGCCCAAGAGCCAGCACTGGACAGGSCCT 7738
Db 646 GCGGAGTCCAAAGCCAGAGAGCGCCCGCGCCCAAGAGCCAGCACTGGACAGGSCCT 705
QY 7739 GGCACCGGCGGAGGTCAGTGTCCAGACCCCGCGGAGGAGCTCTCGGACTGCGGTGG 7798
Db 706 GGCACCGGCGGAGGTCAGTGTCCAGACCCCGCGGAGGAGCTCTCGGACTGCGGTGG 765


```
Qy 7799 CACAGCCGAGGCCCTGAAGGTGCAGACCGGCTCAGTGTGCCAGCACCCGCGAGCATCAC 7858
Db 766 CACAGCCGAGGCCCTGAAGGTGCAGACCGGCTCAGTGTGCCAGCACCCGCGAGCATCAC 825
Qy 7859 ACCGCCGAGGGTCCAGTGTCTACTTCCAGAGTGCAGAGATGAAGGCGCATGAAGGAGCT 7918
Db 826 ACCGCCGAGGGTCCAGTGTCTACTTCCAGAGTGCAGAGATGAAGGCGCATGAAGGAGCT 885
Qy 7919 GCTGTGTGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCGAGGCACACAGTGC 7978
Db 886 GCTGTGTGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCGAGGCACACAGTGC 945
Qy 7979 AGTCAGATGAAGAGCAGAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTCTCTCAA 8038
Db 946 AGTCAGATGAAGAGCAGAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTCTCTCAA 1005
Qy 8039 CGCGAGCGGCAAGCCCTCTCAACTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGC 8098
Db 1006 CGCGAGCGGCAAGCCCTCTCAACTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGC 1065
Qy 8099 TCCG-CCTTAGTCCCCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGC 8157
Db 1066 TCCGCCCTTAGTCCCCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGC 1125
Qy 8158 CTCACCTCTACCCAAACCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCACGA 8217
Db 1126 CTCACCTCTACCCAAACCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCACGA 1185
Qy 8218 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCCAGTGGTTAGATATAGGTCCTCTCTCC 8277
Db 1186 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCCAGTGGTTAGATATAGGTCCTCTCTCC 1245
Qy 8278 CAGAAATCCACGCGCCCAATGGAACCTCAGCGCTGGGTCTCTAAATACAGCTCTTTAAAG 8337
Db 1246 CAGAAATCCACGCGCCCAATGGAACCTCAGCGCTGGGTCTCTAAATACAGCTCTTTAAAG 1305
Qy 8338 CCCAGCCCTTGAAGAACCAAGCTCTCTCTCGGAACCGCTACCTAGAGCCAGACCAAGCT 8397
Db 1306 CCCAGCCCTTGAAGAACCAAGCTCTCTCTCGGAACCGCTACCTAGAGCCAGACCAAGCT 1365
Qy 8398 TACTCAGGCTCTCCCAAGCTGTAGAGAGCTGAGGTTTCACCCCTTAACCCAGGAGCAGC 8457
Db 1366 TACTCAGGCTCTCCCAAGCTGTAGAGAGCTGAGGTTTCACCCCTTAACCCAA-GGAGCAG 1424
Qy 8458 AGGTCCCACTCCAGCCCGGAGGAGCTAGGACCACTCAGCCCTTAGGAGTATATTTCGCG 8517
Db 1425 AGGTCCCACTCCAGCCCGGAGGAGCTAGGACCACTCAGCCCTTAGGAGTATATTTCGCG 1483
Qy 8518 ACTTCAGAAATTCATATCTCGGAATCCCAAGCTCCCTGCCCCCAATACCTCAGTCTCTG 8577
Db 1484 ACTTCAGAAATTCATATCTCGGAATCCCAAGCTCCCTGCCCCCAATACCTCAGTCTCTG 1543
Qy 8578 TTCAGAAATTTGGAATCTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAC 8637
Db 1544 -TCCAGAAATTTGGAATCTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAC 1602
Qy 8638 TCCGCCCCAGCCTATGAGCATCTGAGCCCGGCCCTCTTCTGACGAAATCGCCCGCG 8697
Db 1603 TCCGCCCCAGCCTATGAGCATCTGAGCCCGGCCCTCTTCTGACGAAATCGCCCGCG 1662
Qy 8698 ATCAGAGCAGACCTCTCTCGACCTCTCGGAACCTCCAGAGGTCACGCCATCTCG 8757
Db 1663 ATCAGAGCAGACCTCTCTCGACCTCTCGGAACCTCCAGAGGTCACGCCATCTCG 1722
Qy 8758 GAGCATCCCGAGGAATCTCCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCTT 8817
Db 1723 GAGCATCCCGAGGAATCTCCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCTT 1781
Qy 8818 CTGTCTTTGTACATAGATTTATTTTTCAGTTCCAGAAAGATGAATACATTTTGT 8873
Db 1782 CTGTCTTTGTACATAGATTTATTTTTCAGTTCCAGAAAGATGAATACATTTTGT 1837
```

```
RESULT 15
AF109390
LOCUS Mus musculus menin (Men1) gene, complete cds. 12-JUN-1999
DEFINITION
ACCESSION AF109390
VERSION AF109390.1 GI:5051649
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 9286)
Guru,S.C., Crabtree,J.S., Brown,K.D., Dunn,K.J., Manickam,P.,
Prasad,N.B., Wangsa,D., Burns,A.L., Spiegel,A.M., Marx,S.J.,
Pavan,W.J., Collins,F.S. and Chandrasekharappa,S.C.
Isolation, genomic organization, and expression analysis of men1,
the murine homolog of the MEN1 gene
Mamm. Genome 10 (6), 592-596 (1999)
99272818
PUBMED 10341092
2 (bases 1 to 9286)
Guru,S.C., Crabtree,J.S., Brown,K.D., Dunn,K.J., Manickam,P.,
Prasad,N.B., Wangsa,D., Burns,A.L., Spiegel,A.M., Marx,S.J.,
Pavan,W.J., Collins,F.S. and Chandrasekharappa,S.C.
Direct Submission
Submitted (25-NOV-1998) Genetics and Molecular Biology Branch,
National Human Genome Research Institute, NIH, Bldg49, Rm3C36, 9000
Rockville Pike, Bethesda, MD 20892, USA
FEATURES
Location/Qualifiers
1..9286
/organism="Mus musculus"
/strain="129Sv/cj7"
/db_xref="taxon:10090"
/chromosome="19"
2962..8815
/gene="Men1"
join(2963)..3829,4649..4857,5077..5205,5403..5443,
5525..5612,6207..6343,6750..6885,7263..7427,7611..8096)
/gene="Men1"
/product="menin"
join(2963)..3006,3369..3829,4649..4857,5077..5205,
5403..5443,5525..5612,6207..6343,6750..6885,7263..7427,
7611..8096)
/gene="Men1"
/product="menin"
join(3385)..3829,4649..4857,5077..5205,5403..5443,
5525..5612,6207..6343,6750..6885,7263..7427,7611..8096)
/gene="Men1"
/note="similar to Homo sapiens MENIN"
/codon_start=1
/product="menin"
/protein_id="AAB38333.1"
/db_xref="GI:5051650"
/translation="MGLKAAQKTLPLRSTDDVVRLPAAELGREEDPLVLSLVLGPD
EHLFVNRVPTWVPELTFQSPADPPGGLTFPPVADLSIIAALYARFTAQIRGAVD
LSLPRGGVSSRELKVKSDVWNSLSYFKDRAHQSLFSTITKLDSSGVAER
VGCAGALGRDLVHALSEDHAWFGNGEATVETWKGKNGDRGQTVNAGVAER
SWLYLKSGVMRCRMEVAFVCAINPSIDLHTDSLELLQLOKLLWLLYDIAGHLERY
PMALGNLADLELEPTGPRDPLTLVHKIGASAKTYVDEHLYPYMYLAGYHCRNRV
REALQAWADTATVTDYNYCREDEELYKEFEVANDVLPNLKEASLLETGERTGE
QAGTCQGSALQDPECFAHLRFYDICKWEGSPPLVLIHGWATFVLSLGRFEQ
VRQKHIVISREAEAAAEPEWDEARERRRGRPRRESKPEEPKPPKPLDKGPGSQ
SAGSGPPRKTSGTVPGTGTGGVQEAQAPAAAPASPPPEPVLTTFQSEKMKMKELL
VATKINSSAIKLQLTQASQVQMKKQKYSTPSDVTLSFLKQRKGL"
BASE COUNT 2104 a 2383 c 2518 g 2281 t
ORIGIN
```

```
Query Match 11.7%; Score 1075.8; DB 94; Length 9286;
Best Local Similarity 60.1%; Pred. No. 3,7e-218;
Matches 2957; Conservative 0; Mismatches 1473; Indels 488; Gaps 51;
Qy 4272 CTTTCATTACCTCCCTTCCACAGGCACCAANTGGACAGCTCCGGTCTGCCCTTTCCT 4331
```


Db 4624 CTCTCATGACTGCCCCCTTTACACAGGCACCAAACTGGACAGCTCGGGCGTGGCCTTTGCA 4683
QY 4332 GTGGTTGGGGCCCTCCAGGCCCCCTGGGTCTCCGGGATGTCACCTCGCCCTGTCTGAGGAT 4391
Db 4684 GTGGTAGGGCCCTCCAGGCTCTGGGTCTCAGAGATGTCATCTGGCCCTGTCTGAAGAT 4743
QY 4392 CATGCTCGGGTAGTGTGTTGGGCCAATGCGGAGCAGACAGCTGAGGTCACTGGCAGGC 4451
Db 4744 CATGCTGGGTGGTGTGTTGGGCCAACGGGGAGCAGACAGCTGAGGTGACGTGGCAGGC 4803
QY 4452 AAGGGCAACGAGCAGCCGAGGGGCCAGACAGTCAATGCCGGTGTGGCTGAGCGGGTATTG 4511
Db 4804 AAAGGCAACGAGCAGCCGAGGGGCCAGACAGTCAATGCCGGTGTGGCTGAGCGGGTATTG 4863
QY 4512 TTCCCTCCCCCAGGCTTGTGCCCTTCATAGTGTAGTCCCAAGCCACCAAGGAGACTC 4571
Db 4864 TATCTTCCCTCCA---TTGTGCCCTTTTCCCTTCTGTGAGTCTAAGTGACCTGAGGATTC 4920
QY 4572 CATTTTCTTGGGCCACAC-----CCCTTTCTTCCATCACACCACACATAGG 4618
Db 4921 TGCTTCCCGAGCGAGCTCCCTTACCAGTCTCCCTCTGTGTTTATGTATCCAAAGCGGGG 4980
QY 4619 AAGGGAACAGAGAGCCCTTTTCTGCTGTCTATCTCCCTGAAGCAGCAGGCTGG 4678
Db 4981 GCGAGGCAAGGCGAGGTGAATTTACGTTGCTGTGCCCTAGAGCAGGCTCAAGGTGG 5040
QY 4679 GCATCATGAGACATATGATGATCATCCCCCTTAGAGCTGCTGTACCTGAAAGGATC 4738
Db 5041 TACATCTGAGATGT-CTGCTTTCTCCCTTCCAGAGCTGGCTGTACCTGAAAGGGTC 5099
QY 4739 ATATGCGCTGTGACGCGAAGATGAGGTGGGTTTCATGGTGTGCCATCAACCCCTTC 4798
Db 5100 GTACATGCGCTGCGACCGTAAAGATGAGGTGGGCTTCATGCTGTGCCATCAACCCCTTC 5159
QY 4799 CATTGACCTGCACACGAGCTGCTGAGCTCTTGACAGCTCCAGAGTGGGCTGAGCC 4858
Db 5160 CATCGATCTTCACACTGACTCTTTGGAACCTGTGCAACTACAGCAGGTGAGGACTGCC 5219
QY 4859 AATGGGCGAGGACTGGGCTAGGCCAGACTTGACTTGCTGTGGGACCCCTGGGCGAGGCAC 4918
Db 5220 -----AATGCACACGAAG 5219
QY 4919 TTTCCCTTCTGAGCTTACGTTCCCTCTCGGAAATGGGTAGTAATTCCTGGCCTG 4978
Db 5220 ---CCAGCCCGAGGACCTTCCCTTCCCTTCTGTACAGAGAATGATCCTTTTC- 5269
QY 4979 GCCTTTCCAGGCTCTTGGGAGATAGAATTGAGATGTGAATGCTTTGATCCATTA 5038
Db 5270 -----AATGCACACGAAG 5283
QY 5039 AAGGCTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGTGCCTGTGTTCTGAC 5098
Db 5284 GCCTTTGAGTCTTTACTCTAGCCCTTC--TGTGGTGGGAGGGCGCTGTTCTCTGGT 5341
QY 5099 CCCCACCTCTGCCCGATAGGCTAAGGACCGCTTCTCCCTGCTGTCGGTGGCTCATAACT 5158
Db 5342 CACTTCTAGCTAGCTGGG-----GCTCCCTTCTACTTACCAAGCC 5384
QY 5159 CTCTCTTCCGCTCTAGAGCTGCTCTGCTGCTATGACCTGGGACATCTGGAAGG 5218
Db 5385 CTCTCCCTTTGGCTCTAGAGCTGCTCTGCTGCTGTATGACCTCGGACATCTGGAAGG 5444
QY 5219 TCAGTAGAGGAAGTGGCCAGGCTGCGCTGTGAGCGCGGGGCTGGTGGCAGCCTG 5278
Db 5445 TCAGTGAAGGCTATGATACAGATAGGCTTTGTGGGATCAGGAACTGAGTATAGCCTG 5504
QY 5279 A-ATTATGATCTTTCTAGGTACCCCATGCCCTTAGGAACTCTAGGAGCTCTAGAGGAG 5337
Db 5505 ATTTTATGATCTGTCCAGATACCCCATGCCCTAGGAACTTGGCAGACCTTGGAGGAG 5564
QY 5338 CTGAGCCCAACCTTGGCGGCGAGCCCACTACCCCTCTACCAAGAGGTGGGGGCACT 5397
Db 5397 -----AATGCACACGAAG 5283

Db 5565 CTGAGGCTTACCCCGGCGCGGCGAGACCCCACTCACCCTTTATCACAAGGTAAGGGCATCT 5624
QY 5398 AAGGAGGCTCAGAAGGAGAGACCTTAACAGTGGCTAGGCGAGGGGCTCATCTGGGCAG 5457
Db 5625 CA-----GAAACAGCAGCTCTTCTCTGGGATGGAGAGGAG-----AGAGGTG 5668
QY 5458 ATGAGAAAGAACTTTGTGTGTTGGGGGATATCGCCCATCCAGTCTCAGTTGTGTCAAC 5517
Db 5669 AAAAGGGAAGAACTCTGTGTGTTGGAGGTAT-----TTCTGCTTACTTTGTGTAGGG 5723
QY 5518 TGTGTGCAGAACTCAGTTCACTCAGGCTGTCTGAGGGGTCTGAGGGGTCCAGGGTCCCGAGCCTGG 5577
Db 5724 CATGTGTGTAATGAGGTAGTTCTGTGTTTAAAGAGAG---CCCTGGCTCCCCAGTAG 5780
QY 5578 GAGTGGCAGGGGTGCAATTTGTCCCTCAGCCCTGCTTTTCTGCCACTGCTTACTCTCC 5637
Db 5781 GAGTGGCAGGGCCGATTTGATCCCAAGCCCTGCAGTTTTCAGTCCCTTTATTATCC 5840
QY 5638 TTCTGTGAGTATTAACAGAGGTCAAAATGTGTAGAGCAGCTGAAGAGGGGTGTTCACATTG 5697
Db 5841 TGGCTATAGCTCTGTGGTAG---CACTGTGGAGAGATGTTAGCTTGTCTCATTTGTTAGGA 5897
QY 5698 GTGGGTGAGTGGGAGGAGGGCCATTGGCTGGGCTTGAAGTCTTTTGGTATGTGTA 5757
Db 5898 GGCTGGTTGGGCTTGAAGTCTTTTACAGAGACCATGTGGTGAAGAACTGCCCCAGGCT 5957
QY 5758 GAAGAGTCTCTGAGAAAGAGAGGGCCCTGAGCTCGGAGGGCAGGCCCCCCTGCGACT 5817
Db 5958 TGAGTGTTCCTGCTCCTCCTAGGCTTGTGTTGGGGT---CACCACTCCTACTGT 6014
QY 5818 CTGCCCGAGGCTCAGCAGCAGTCTCTAGACCCAGGGAGGAGACCAAGGTAGAAAGGCT 5877
Db 6015 CTG-CCCTTCTTTAGCAAGCAGGCTTGTAGTCCAGGCTCGAATGAGGTGGAGGGA 6073
QY 5878 GGCAGCAGTGGAGTGGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCTGTGGCC 5937
Db 6074 GTATGGAACACCAAGG--CAAGGGGCGCAGCAGTATTTAGAGACAGGGAGGACCTTTAGCAT 6132
QY 5938 CTCTGCTGCTGCTGCTGCTGGGCAATTTGTGCCAGCAGGCGAGCTGGGGTGCCTCCCTG 5997
Db 6133 CCTTTTCTGT-----AGGCTATCAGTGTAGGAGCAGCTGGGGCTGCTCCTG 6182
QY 5998 AGGATCTCTGCTCAGTCCATCCAGGCAATGCTCTCAGCAGCAAGACTACTATCGGAT 6057
Db 6183 AGGTCTCTCTG---CTCCATCTCTCAGGAAATGCTCAGTAAAGACTACTACCAAGAT 6239
QY 6058 GAACACATCTACCCCTACATGTACCTGGCTGGCTTACCACTGTCCCAACCCCAATGTGGG 6117
Db 6240 GAACACATCTACCCCTACATGTACCTGGCTGGCTTACCACTGTCCCAACCCCAATGTGGG 6299
QY 6118 GAAGCCCTGAGGCTGGGCGGACACGGCCACTGTCTATCCAGGAGTGAAGTATCCCTAC 6177
Db 6300 GAGCCCTGAGGCTGGGCGGACACTGCCCCATGTTATCCAGAGTGAAGATTCCTCTGC 6359
QY 6178 TAGGGCTGAGGCTGTCTTCTTCCCTCCATCAGTGTTCACACCAACCCCTGCTCAGGA 6237
Db 6360 CAGGG-TTCCAGGCTGCCCTTGT---CTCTATGACTTTTCCAA--TCGGTCACTCCAGTT 6412
QY 6238 CTGAGGCTGGCTCCAGCGCCCATCCCTTCCATCCAGTCCCTAGGCAGCAAGGCCAC 6297
Db 6413 GTGAGACTTGTCTCAGTCTGTCTCTCTGTATCAGGCCCCCTATACAG---TCAAGGCTAC 6469
QY 6298 CATTAACAGGAGTGAAGGACCCCTGATTAAAGGTGTACATCTTTCCCTGCCCTCCCTCTC 6357
Db 6470 CATTAATGGGAAGTGAAGACCCCTGGTGTAGTCTGCTCTCTCTCATGCTTCTGTCA 6529
QY 6358 CTCTAAATTTTCTCAGAAAGTCTCAAAATCTCCAAATGTTTAAACCACCAATCAACC 6417
Db 6530 CTCTCTCTGCCAGTTTCTCTCTTATAGTCAAGTGTGTTATCCATGGTGTCTGT 6589
QY 6418 AGCAGTGGGACTTCCACCCCTCGGCCCATCCCTCTCTCTCTCTCTCTCTCTCTCTCT 6477
Db 6590 CTCTGCCCAACTTCTAGTCCAGTCCCAACCCACATCCCACTTCCCTTGTGACTAGC 6649

Qy 6478 GGGCTGACCCAGACAGCATCATTTTTCAGTGAAGACCCCACTACTCCCCAGCCCTTG 6537
Dy 6650 ATGATCACATCTGGGTGAACA-----AGGACTCCACCTCGCACCCAGTCACTGG 6699
Qy 6538 GGGCTCCATCCCGCCAGTGCCTTGGGGTACCCCGAGTGGTGAAGACCCCTTCAAGACC 6597
Dy 6700 G-----ATGTCCCATTTGTGGGCAACT-----GTC 6724
Qy 6598 TACAGAGACCCCACTGCTCTCACAGCTACAACTACTCGCGGAAGACGAGAGATCTACA 6657
Dy 6725 AACTGAATCTGTGTGTCTCACAGCTACAACTACTCGCGGAGATGAGAGATCTACA 6784
Qy 6658 AGGAGTTCTTTGAAGTACCAATGATGTATCTCCCAAGCTCTGAAGGAGGACGACGCT 6717
Dy 6785 AGGAATCTTTGAAGTGCCCAATGACGTATCCCAAGCTCTGAAGGAGGCGCCAGCC 6844
Qy 6718 TGCTGGAGCGGGAGAGAGCGGGGGAGCAAGCCAGGTGAAGGCTTGAGACTCCA 6777
Dy 6845 TGCTAGAAACAGCGAAGAGCGGACTGGGGAGCAAGCCAGGTGAAGGGCTGG-----CCT 6900
Qy 6778 GCCTGTGTCCAGCTCCCACTGACAGGGCTCCCTTCCACAGGGCCATGGGGCTGCAT 6837
Dy 6901 CCATTGGCGCACCTCCCTCGCGGCA-----CATGTATGAGACCGCAA 6944
Qy 6838 GTACGGGATTAGGATGCGAGGAGGAAGGTGGCGGACTGAAGGTATTTTAGAGTTTCTA 6957
Dy 6945 GGACAGGATTGGGAAT---TCTCTTTTGTGAGCTGATCCTCGGCGAGGGTTTGAAG 7001
Qy 6898 GCTATACTAGGTCTGGGCGCCAGGTGGACGGGACTGAAGGTATTTTAGAGTTTCTA 6957
Dy 7002 GATTTAGGGGTTTATAGACTAAGCTTTTAGAGAAATGGG---TTGATGGTCCATCT 7057
Qy 6958 CCCTGTGCTTCAGTTTCATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGTAGGGA 7017
Dy 7058 CTCGTAGGTGAGTTCTAAGCTGGCTAGGCTGAGTGAAGAAAGGCGCCCTCTCT 7117
Qy 7018 TGGTACGTCCTGGCTATGGATTTGCTTTTATAAAGAAAGAGTTCTAAGAAATGTTCCCA 7077
Dy 7118 CCCCTCACTCGCACTAAGGGTTGGTTTATATAAAGAAAGAGGT-----CCT 7164
Qy 7078 ACCTATGCTTACCTTTCTGAGCCAGGGGTCTTTGCTAGTGGGGGCTGCGCTGTG 7137
Dy 7165 GCCTGCAATTCCTCTCCCAAGCTGAGGCTTTGCCAGGTGGAGGAGCGCTGGCCGCA 7224
Qy 7138 CCCTCTGCTAAGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTCGTCCAGGG 7197
Dy 7225 CGCTCTG-----TGACCTCTGCTGTAACCTTCCCTCTATATCCAGGG 7264
Qy 7198 CACCCAGACCAAGTTCCGCCCTCCAGGACCCCTGAGTGTTCGCCCACTGCTGCGGAT 7257
Dy 7265 CACCCAGGCGCAAGTTCTGCTCTCCAGGACCCAGAAATGCTTTGCCCACTGCTGCGGAT 7324
Qy 7258 CTAGCAGGCACTGCAAAATGGAGGAGGCGAGTCCCAAGCTGTGCTGACGTTGGGCTG 7317
Dy 7325 TTATGATGGCATCTGCAAAATGGAGGAGGGTAGCCCCACACGACTGACGTTAGGCTG 7384
Qy 7318 GGCACACCTTTCTGTGCACTCCCTAGGCGGTTTTGAGGACAGGTGAGGACAGCTGCAC 7377
Dy 7385 GGCACATTTCTGCTGTGAGTCCCTAGGCGGCTTCGAGGACAGGTGAGAGGGCCATACC 7444
Qy 7378 AGAGTCTGGGCACTACAGGTGGTGACAGCAAGCCAGCGGGCTTGTACAGATTTTCTGGGCC 7437
Dy 7445 AG-GGCATGGGAACCTACAGGTGGTGACAGTAGTTGTAGGGGAGGAGGTATCATTTGCATG 7503
Qy 7438 AGGGCAGCATCTGCCATCCCTTCCGTGCGGATGGGACTGAGACCCCTGGGTGGAT 7497
Dy 7504 TGTATGAGTA-----CAGATGAGGAGGACACTGCTA-----GA 7535
Qy 7498 GGGATGGCCAGGAGGCTCCTGGAGTTCCAGCCACTGGCGGCAACCTGCTCTCACT 7557
Dy 7536 GGGGTGGCCAGGAGTGAAGCTAGAGGTTCCACAGGATGATCTCAGCCCTTTCACCTTGCT 7595

Qy 7558 TGCTCTCCCACTGGCCCAAGTGGCGAGAAAGTGGCATATAGTGAAGCCGAGAGCCGAGG 7617
Dy 7596 TGCCT-----TTTGCTCAGGTGGCGAGAAAGTGGCATATAGTGGCGGGAAGCAGAGG 7650
Qy 7618 CGGCGGAGGCGGAGGCGGTGGGGCGAGGAAGCCCGGAAGGCGGGCGGGGCGCCAC 7677
Dy 7651 CGGCGGAGGCTGAAGAACCATTTGGGGGATGAAGCCGGAAGAGCGGTGGGTGTCCTCC 7710
Qy 7678 GCGGGAGTTCACCCAGAGAGAGCCCGCGCCCAAGAACCCAGCATGTGACAAAGGCGC 7737
Dy 7711 GAAGAGAGTCCACAGCTGAGGAGCCACACCCCAAGAGCCCTGCAATGGACAAAGGCC 7770
Qy 7738 TGGCACCGGCGAGGGTGCAGTGTGAGACCCCGGGAAGCCCTCTGGGACTGTGCGTG 7797
Dy 7771 CGGCTCAGGACAAAGTGCAGGGTCGGGACCACTAGGAAAACGTGAGGACTGTCCAG 7830
Qy 7798 GCACAGCCC---GAGGGCCTGAAGTGGCAGCACGGTGCAGTGCACGACCCGACCAT 7854
Dy 7831 GTACTACCCCGGTGGCGCAGGAAGTGGCAATGTGTCTCAGGCTCCAGCAGCTGCAGAT 7890
Qy 7855 CACCACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGCGCATGAAGG 7914
Dy 7891 CGCCACCCAGAGGGCCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGATGAAGG 7950
Qy 7915 AGTGTGTTGGCCACCAAGATCAACTGCGAGGCGCATCAAGCTGCACTCAAGGACAGT 7974
Dy 7951 AGTACTGTTGGCCACCAAGATCAACTGCGAGGCGCATCAAGCTGCACTCAAGGACAGT 8010
Qy 7975 CGAAGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGAAGTACACTGTCTTTCC 8034
Dy 8011 CGAAGTGCAGATGAAGAGCAGAAAGTGTCCACCCAGGAGCTACACACTCTCTTTCC 8070
Qy 8035 TCAAGCGCAGCGAAAGGCTCTGAACACTGCGGAGCTTCGGACCGCTTGTGGGAGCC 8094
Dy 8071 TAAAGCGCAGCGAAGGGCTCTGA-----ACTCTGATTCGTGCGCAAGGCC 8120
Qy 8095 AGGCTCGCCTTATTCGCCCACTCTGAGGCCATGTCTGCCCGCCAGCCCAAGAGGAGCA 8154
Dy 8121 AGGCTGCGCTCACTCCAGCAGGCTTCACCCCTCATCTCTTAATTCAGGCTA-GACTCC 8179
Qy 8155 GGCCTCACTCTACCCAAACCTAGTTTCCGGTCCCGGTACAGTCTGTATCAACCCCA 8214
Dy 8180 GCTCTAGTCTAGGGACCCAGAGCTGTCTGAGTACTA-----GGAGCCTT 8229
Qy 8215 CGATTTTCTCCAGCTGAGAACCCAGGCTCTGCCAGTCTGTAGAAATATAGTCTCTTC 8274
Dy 8230 CATTTAGCCCTCTTCAGGCGCCAGAGCCCTGGCTCAGTCTTGAATGTAGGTCTCATC 8289
Qy 8275 TCCAGAGTCCAGCGGCCAATGAAACCTCAGCTGGGTCTTAATTTACAGCTTTTAA 8334
Dy 8290 CCC-----ACGAGCGGCCAATTTGGAACCTCATTTAGGTCCAATCCCGGAAATTCAG 8343
Qy 8335 AGGCCAGCCCTAGAAACCCAAAGCTCCTCTCGGAACCGCTCACTAGACGACGACAA 8394
Dy 8344 TGGACCATGAC--AGGAGCCCAAGTACGCTGAAACAGTAGGCAATTCCTCCCCTTGGG 8401
Qy 8395 CGTTACTCAGGGCTCTCCCAAGCTTGTAGGAGCTGAGGTTTCAACCTTAAACCCAGGAG 8454
Dy 8402 GGCTAGATTGACCTTAGCCC-----CAGAG 8426
Qy 8455 CACAGTCCCACTCCAGCCCGGGAGGCTAGGACCACTCAGCCCTTAGGAGTATATTTC 8514
Dy 8427 CACAGTTGCTACCT-----GGAAACCAAGAGCACTTCACTGCTTCCCTAGGCTGAGTTTC 8478
Qy 8515 CGCACTT---CAGAAATTCATATCTTCGGAATCAAGCTCCTGCCCAATTAACCTCA 8570
Dy 8479 CGCACTTTATATAGATTCCCATCTTTGGGAATCCAAATCCCTTACCCCAAGAACATTAC 8538
Qy 8571 GTCCTGCTTCCAGAAATTTGGAATCTAGTTTCTCTCTCTCTGATCCCTGAGTCTGGGAC 8630
Dy 8539 ATATTGCTGCCGAATTTGGAATC-----TTCTCTGCTCTCTCTAGCCCTTGAAGTACAGG 8593
Qy 8631 ACAAACTCCCGCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTGACGCAAACTG 8690

[illegible]

Search completed: October 14, 2001, 04:12:45
Job time: 29981 sec

THIS PAGE BLANK (USPTO)